

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 05:35:31 ; Search time 172 Seconds
(without alignments)
1072.585 Million cell updates/sec

Title: US-09-886-954A-1
Perfect score: 2457
Sequence: 1 MTPEDPEEQPLLPGGSA.....CVPETKGTLEQITAHFEGR 477

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues
Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A_Geneseq_16dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Table with columns: Result No., Score, Match, Length, DB ID, Description. Contains 25 rows of search results.

Table with columns: ID, Accession, Score, Length, Description. Lists various protein entries and their scores.

ALIGNMENTS

RESULT 1
AAE04888
ID AAE04888 standard; protein; 477 AA.
AC AAE04888;
XX
DT 10-SEP-2001 (first entry)
XX
DE Human transporter and ion channel-1 (TRICH-1) protein.
XX
KW Human; transporter and ion channel-1; TRICH-1; vaccine; cystic fibrosis; gene therapy; amyotrophic lateral sclerosis; amnesia; muscular dystrophy; hypertension; angina; neurological disorder; asthma; bipolar disorder; dementia; depression; Alzheimer's disease; epilepsy; mood; arrhythmia; Pick's disease; ischaemic cerebrovascular disease; AIDS; anxiety; stroke; Huntington's disease; Parkinson's disease; cerebral neoplasm; allergy; demyelinating disease; mental disorder; Schizophrenia; polymyositis; muscle disorder; cardiomyopathy; cataract; myocardiitis; Grave's disease; dermatomyositis; diabetes mellitus; immunological disorder; psoriasis; rheumatoid arthritis; Sjogren's syndrome; systemic lupus erythematosus; sickle cell anaemia; Wilson's disease; infertility; Cushing's disease; scleroderma; pulmonary artery stenosis; neutropenic; Addison's disease; malabsorption syndrome; hypercholesterolaemia; cancer.

Key Domain Location/Qualifiers
Domain /note= "Sugar transporter domain"
Domain /label= Transmembrane_domain
Domain /label= Transmembrane_domain
Domain /label= Transmembrane_domain
Domain /label= Transmembrane_domain
WO200146258-A2.
28-JUN-2001.
22-DEC-2000; 2000WO-US035095.
23-DEC-1999; 99US-0172000P.
14-JAN-2000; 2000US-0176083P.
21-JAN-2000; 2000US-0177332P.
28-JAN-2000; 2000US-0178572P.

CC The present sequence is that of human solute carrier type 2A (SLC2A)
 CC polypeptides 1273153 and 7657681. In the present invention, genetic
 CC screens were designed to identify modifiers of the p53 pathway in
 CC Drosophila in which p53 was overexpressed. Human orthologues
 CC (polynucleotides and polypeptides) of one such modifier were then
 CC identified, including the present polypeptides. SLC2As are glucose
 CC transporter proteins with sugar transporter domains. They are attractive
 CC drug targets for treatment of pathologies associated with a defective p53
 CC signalling pathway, such as cancer. The invention provides in vitro and
 CC in vivo methods of assessing SLC2A function. Modulation of an SLC2A or
 CC its binding partners is useful for understanding the association of the
 CC p53 pathway and its members in normal and disease conditions and for
 CC developing diagnostic and therapeutic modalities of p53-related
 CC pathologies; SLC2A-modulating agents that act by inhibiting or enhancing
 CC SLC2A expression, directly or indirectly, e.g. by affecting an SLC2A
 CC function such as transport or binding activity, can be identified using
 CC methods provided. Modulators include small molecules, nucleic acids,
 CC antibodies, antisense oligonucleotides and phosphothioate morpholino
 CC oligomers (claimed)

SQ Sequence 477 AA;

Query Match 100.0%; Score 2457; DB 6; Length 477;
 Best Local Similarity 100.0%; Pred. No. 8.8e-227;
 Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTPPEEETQPLLGPPGSGAPRRVFLAAFAAALGPLSFGFALGVSSPAISLQRAAPP 60
 DB 1 MTPPEEETQPLLGPPGSGAPRRVFLAAFAAALGPLSFGFALGVSSPAISLQRAAPP 60
 QY 61 APRLDDAAASWFGAVVTLGAAAGVVLGAAAGVLDVDRAGKLSLLCSVFFVAGFAVITAAQDV 120
 DB 61 APRLDDAAASWFGAVVTLGAAAGVVLGAAAGVLDVDRAGKLSLLCSVFFVAGFAVITAAQDV 120
 QY 121 WMLLGGRLLTGLACGVASLAVPVIYSEIAYPAVRGLLSCVQLMVMVVGILLAYLAGWVLE 180
 DB 121 WMLLGGRLLTGLACGVASLAVPVIYSEIAYPAVRGLLSCVQLMVMVVGILLAYLAGWVLE 180
 QY 181 WRWLAVLGCVPSPSLMLLMLCFMPEPRFLLTQHRROEAMALRFLWGSQGWEDPPIGAE 240
 DB 181 WRWLAVLGCVPSPSLMLLMLCFMPEPRFLLTQHRROEAMALRFLWGSQGWEDPPIGAE 240
 QY 241 QSFHLALLRQPGIYKPFIIIGVSLMAFQQLSGVNAVYFAETIPEAKFKDSSLASVVVGV 300
 DB 241 QSFHLALLRQPGIYKPFIIIGVSLMAFQQLSGVNAVYFAETIPEAKFKDSSLASVVVGV 300
 QY 301 IQVLFVAALIMDRAGRELLLVLSGVVMVFTSAPGAYFKLTQGGPGNSSHVAISAPVS 360
 DB 301 IQVLFVAALIMDRAGRELLLVLSGVVMVFTSAPGAYFKLTQGGPGNSSHVAISAPVS 360
 QY 361 AQPVDASVGLAWLAVGSMCLFIAGFVAVGMPPIWLLMSEIFPLHVKGVATGICVLTNWLML 420
 DB 361 AQPVDASVGLAWLAVGSMCLFIAGFVAVGMPPIWLLMSEIFPLHVKGVATGICVLTNWLML 420
 QY 421 AFLVTKFESSLMEVLRPYGAFWLASAFICFVFLTLFCVPEPKTKLEQITAHFEGR 477
 DB 421 AFLVTKFESSLMEVLRPYGAFWLASAFICFVFLTLFCVPEPKTKLEQITAHFEGR 477

RESULT 3
 AAB66932
 ID AAB66932 standard; protein; 477 AA.
 XX
 AC AAB66932;
 XX
 DT 17-APR-2001 (first entry)
 XX
 DE Human GLUTX1.
 XX
 KW Human; GLUTX; gene therapy; vaccine; hexose transport modulator;
 KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia;
 KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 XX

OS Homo sapiens.
 XX WO200104145-A2.
 XX 18-JAN-2001.
 XX 14-JUL-2000; 2000WO-IB001042.
 XX 14-JUL-1999; 99US-0143907P.
 XX 27-AUG-1999; 99US-0151140P.
 XX 23-FEB-2000; 2000US-0184285P.
 XX 13-JUL-2000; 2000US-00616132.
 XX (UYLA-) UNIV LAUSANNE.
 XX Thorens B, Ibberson M, Uldry M;
 DR WPI; 2001-112615/12.
 DR N-PSDB; AAF55865.

PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in the
 PT prevention, diagnosis and treatment of hexose transport disorders, e.g.
 PT ischaemia and diabetes.
 XX Claim 11; Page 70-71; 124pp; English.

CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facultative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is human GLUTX1
 XX Sequence 477 AA;

Query Match 99.6%; Score 2448; DB 4; Length 477;
 Best Local Similarity 99.6%; Pred. No. 6.4e-226;
 Matches 475; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTPPEEETQPLLGPPGSGAPRRVFLAAFAAALGPLSFGFALGVSSPAISLQRAAPP 60
 DB 1 MTPPEEETQPLLGPPGSGAPRRVFLAAFAAALGPLSFGFALGVSSPAISLQRAAPP 60
 QY 61 APRLDDAAASWFGAVVTLGAAAGVVLGAAAGVLDVDRAGKLSLLCSVFFVAGFAVITAAQDV 120
 DB 61 APRLDDAAASWFGAVVTLGAAAGVVLGAAAGVLDVDRAGKLSLLCSVFFVAGFAVITAAQDV 120
 QY 121 WMLLGGRLLTGLACGVASLAVPVIYSEIAYPAVRGLLSCVQLMVMVVGILLAYLAGWVLE 180
 DB 121 WMLLGGRLLTGLACGVASLAVPVIYSEIAYPAVRGLLSCVQLMVMVVGILLAYLAGWVLE 180
 QY 181 WRWLAVLGCVPSPSLMLLMLCFMPEPRFLLTQHRROEAMALRFLWGSQGWEDPPIGAE 240
 DB 181 WRWLAVLGCVPSPSLMLLMLCFMPEPRFLLTQHRROEAMALRFLWGSQGWEDPPIGAE 240
 QY 241 QSFHLALLRQPGIYKPFIIIGVSLMAFQQLSGVNAVYFAETIPEAKFKDSSLASVVVGV 300
 DB 241 QSFHLALLRQPGIYKPFIIIGVSLMAFQQLSGVNAVYFAETIPEAKFKDSSLASVVVGV 300
 QY 301 IQVLFVAALIMDRAGRELLLVLSGVVMVFTSAPGAYFKLTQGGPGNSSHVAISAPVS 360
 DB 301 IQVLFVAALIMDRAGRELLLVLSGVVMVFTSAPGAYFKLTQGGPGNSSHVAISAPVS 360
 QY 361 AQPVDASVGLAWLAVGSMCLFIAGFVAVGMPPIWLLMSEIFPLHVKGVATGICVLTNWLML 420
 DB 361 AQPVDASVGLAWLAVGSMCLFIAGFVAVGMPPIWLLMSEIFPLHVKGVATGICVLTNWLML 420
 QY 421 AFLVTKFESSLMEVLRPYGAFWLASAFICFVFLTLFCVPEPKTKLEQITAHFEGR 477
 DB 421 AFLVTKFESSLMEVLRPYGAFWLASAFICFVFLTLFCVPEPKTKLEQITAHFEGR 477

Db 241 HQGFQALLRRPGIYKPLIIGISLWVFOQLSGVNAIMFYANTIFEBAKFKDSSLASVTVG 300
 Qy 300 VIQVLFATAALIMDRAGRELLLVLSGVVWVFSFARFAYFKLTQSGPGNSSHVAISAPV 359
 Db 301 IIQVLFATAALIMDRAGRELLLVLSGVVWVFSFARFAYFKLTQSGPGNSSHVGIILVPI 360
 Qy 360 SAQPVDAASVGLAWLAVGSMCLFIAGFAVGVGPIPWLLMSEIFFPLHVKGVAATGICVLTNWL 419
 Db 361 SAEPVDSVGLAWLAVGSMCLFIAGFAVGVGPIPWLLMSEIFFPLHVKGVAATGICVLTNWF 420
 Qy 420 MAFLVTKFESLMEVLPYGAFWLASFIPSVLFTLFCVPEPKTKLEQITAHFEGR 477
 Db 421 MAFLVTKFESLMEVLPYGAFWLTAACILSVLFTLFCVPEPKTKLEQITAHFEGR 478
 RESULT 5
 AAB66933 standard; protein; 478 AA.
 XX AAB66933;
 AC AAB66933;
 DT 17-APR-2001 (first entry)
 DE Rat GLUTX1.
 KW Rat; GLUTX; gene therapy; vaccine; hexose transport modulator;
 KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia;
 KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 OS Rattus sp.
 XX WO200104145-A2.
 PN 18-JAN-2001.
 PD 14-JUL-2000; 2000WO-IB001042.
 PF 14-JUL-1999; 99US-0143907P.
 PR 27-AUG-1999; 99US-0151140P.
 PR 23-FEB-2000; 2000US-0184285P.
 PR 13-JUL-2000; 2000US-00616132.
 XX (UYLA-) UNIV LAUSANNE.
 PA Thorens B, Ibberson M, Uldry M;
 PI WPI; 2001-112615/12.
 XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in the prevention, diagnosis and treatment of hexose transport disorders, e.g. ischemia and diabetes.
 PS Claim 11; Page 74-75; 124pp; English.
 CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and AAB66932-AAB66941). The GLUTX proteins are related to the facultative glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport function. The GLUTX proteins may be used in the diagnosis, prevention and treatment of hexose transport disorders such as ischaemia, diabetes, hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a neurodegenerative disease. The present sequence is a consensus sequence for GLUTX1
 SQ Sequence 478 AA;
 Query Match 90.5%; Score 2224.5; DB 4; Length 478;
 Best Local Similarity 88.9%; Pred. No. 1.9e-204;
 Matches 425; Conservative 23; Mismatches 29; Indels 1; Gaps 1;
 Qy 1 MTPEDPEETQPLLPGGGAPRRVFLAAFAAALGFLGFGALGYSSPAIFSLQRAAPP 60
 Db 1 MSPEDEQETQPLRPPGARAPRRVFLAAFAAALGFLGFGALGYSSPAIFSLRRTAPP 60
 Qy 61 APRLDDAASWFGAVVTLGAAGVVLGGWMLVDRAGKLSLLCSVPFVAGFAVITAAQDV 120
 Db 61 ALRLGDAASWFGAVVTLGAAGVVLGGWMLVDRAGKLSLLCTVFPFVITAAQDV 120
 Qy 121 WMLLGGRLTGLACGVASLIVAPYIIEIAYPAVRGLLGSQVLMVVVHLLVLAGWVLE 180
 Db 121 WMLLGGRLTGLACGVASLIVAPYIIEIAYPAVRGLLGSQVLMVVVHLLVLAGWVLE 180
 Qy 181 WRWLVAGCVPPSMLLMLCMCFMPEPFRFLLTQHRQEAAMALRFLWGSQGWEDDPIGAE 240
 Db 181 WRWLVAGCVPPSMLLMLCMCFMPEPFRFLLTQHRQEAAMALRFLWGSQGWEDDPIGAE 240
 Qy 241 -QSFHLALLRQPGIYKPLIIGISLWVFOQLSGVNAIMFYANTIFEBAKFKDSSLASVTVG 299

Db 241 HQGFQALLRRPGIYKPLIIGISLWVFOQLSGVNAIMFYANTIFEBAKFKDSSLASVTVG 300
 Qy 300 VIQVLFATAALIMDRAGRELLLVLSGVVWVFSFARFAYFKLTQSGPGNSSHVAISAPV 359
 Db 301 IIQVLFATAALIMDRAGRELLLVLSGVVWVFSFARFAYFKLTQSGPGNSSHVGIILVPI 360
 Qy 360 SAQPVDAASVGLAWLAVGSMCLFIAGFAVGVGPIPWLLMSEIFFPLHVKGVAATGICVLTNWL 419
 Db 361 SAEPVDSVGLAWLAVGSMCLFIAGFAVGVGPIPWLLMSEIFFPLHVKGVAATGICVLTNWF 420
 Qy 420 MAFLVTKFESLMEVLPYGAFWLASFIPSVLFTLFCVPEPKTKLEQITAHFEGR 477
 Db 421 MAFLVTKFESLMEVLPYGAFWLTAACILSVLFTLFCVPEPKTKLEQITAHFEGR 478
 RESULT 5
 AAB66933 standard; protein; 478 AA.
 XX AAB66933;
 AC AAB66933;
 DT 17-APR-2001 (first entry)
 DE Rat GLUTX1.
 KW Rat; GLUTX; gene therapy; vaccine; hexose transport modulator;
 KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia;
 KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 OS Homo sapiens.
 OS Mus sp.
 OS Rattus sp.
 XX WO200104145-A2.
 PN 18-JAN-2001.
 PD 14-JUL-2000; 2000WO-IB001042.
 PF 14-JUL-1999; 99US-0143907P.
 PR 27-AUG-1999; 99US-0151140P.
 PR 23-FEB-2000; 2000US-0184285P.
 PR 13-JUL-2000; 2000US-00616132.
 XX (UYLA-) UNIV LAUSANNE.
 PA Thorens B, Ibberson M, Uldry M;
 PI WPI; 2001-112615/12.
 XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in the prevention, diagnosis and treatment of hexose transport disorders, e.g. ischemia and diabetes.
 PS Claim 11; Page 74-75; 124pp; English.
 CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and AAB66932-AAB66941). The GLUTX proteins are related to the facultative glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport function. The GLUTX proteins may be used in the diagnosis, prevention and treatment of hexose transport disorders such as ischaemia, diabetes, hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a neurodegenerative disease. The present sequence is a consensus sequence for GLUTX1
 SQ Sequence 478 AA;
 Query Match 90.5%; Score 2224.5; DB 4; Length 478;
 Best Local Similarity 88.9%; Pred. No. 1.9e-204;
 Matches 425; Conservative 23; Mismatches 29; Indels 1; Gaps 1;
 Qy 1 MTPEDPEETQPLLPGGGAPRRVFLAAFAAALGFLGFGALGYSSPAIFSLQRAAPP 60
 Db 1 MSPEDEQETQPLRPPGARAPRRVFLAAFAAALGFLGFGALGYSSPAIFSLRRTAPP 60
 Qy 61 APRLDDAASWFGAVVTLGAAGVVLGGWMLVDRAGKLSLLCSVPFVAGFAVITAAQDV 120
 Db 61 ALRLGDAASWFGAVVTLGAAGVVLGGWMLVDRAGKLSLLCTVFPFVITAAQDV 120
 Qy 121 WMLLGGRLTGLACGVASLIVAPYIIEIAYPAVRGLLGSQVLMVVVHLLVLAGWVLE 180
 Db 121 WMLLGGRLTGLACGVASLIVAPYIIEIAYPAVRGLLGSQVLMVVVHLLVLAGWVLE 180
 Qy 181 WRWLVAGCVPPSMLLMLCMCFMPEPFRFLLTQHRQEAAMALRFLWGSQGWEDDPIGAE 240
 Db 181 WRWLVAGCVPPSMLLMLCMCFMPEPFRFLLTQHRQEAAMALRFLWGSQGWEDDPIGAE 240
 Qy 241 -QSFHLALLRQPGIYKPLIIGISLWVFOQLSGVNAIMFYANTIFEBAKFKDSSLASVTVG 299

CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is murine GLUTX1
 XX Sequence 477 AA;

Query Match 87.5%; Score 2150; DB 4; Length 477;
 Best Local Similarity 85.6%; Pred. No. 2.6e-197;
 Matches 409; Conservative 33; Mismatches 34; Indels 2; Gaps 2;

QY 1 MTPEDPEETPTLGLPGGSGAPRRVFLAFAAALGPLSGFALGHVSSPAISLQRAAPP 60
 Db 1 MSPEDPEETPTLGLPGGSGAPRRVFLAFAAALGPLSGFALGHVSSPAISLQRAAPP 60
 QY 61 APRLDDAAASWFGAVVTLGAAAGVGLGWLDVDRAGKLSLLLSGVVFPVAGFAVITAAQDV 120
 Db 61 ALRLGDTAASWFGAVVTLGAAAGVGLGWLDVDRAGKLSLLLSGVVFPVAGFAVITAAQDV 120
 QY 121 WMLLGGRLLTGLACGVSALVAPVYIASEIAYPAVRGLLGSVCVQLMVTGILLAYVAGWVLE 180
 Db 121 WMLLGGRLLTGLACGVSALVAPVYIASEIAYPAVRGLLGSVCVQLMVTGILLAYVAGWVLE 180
 QY 181 WRWLAVLGCVPVPSLMLLMCFMPEPRFLLTQHRROEAMAAALRFLMGSEGGWEDPPIGAE 240
 Db 181 WRWLAVLGCVPVPSLMLLMCFMPEPRFLLTQHRROEAMAAALRFLMGSEGGWEDPPIGAE 240
 QY 241 -QSFLHALLRQPGIYKPFITIGVSLMAFOQLSGVNAVMFYAEITFEBAKFKDSSLASVVVG 299
 Db 241 HQGFQALMLRRPGVHKPLIIGICLMVYFQQLSGVNAIMFYANTIFEBAKFKDSSLASVVVG 299
 QY 300 VIQVLFATAALIMDRAGRRLLLVLSGVVMVFTSAFGAYFKLITQGGPNSSHVAISAPV 359
 Db 301 IIQVLFATAALIMDRAGRRLLLVLSGVVMVFTSAFGAYFKLITQGGPNSSHVAISAPV 359
 QY 360 SAQPVDAVGLAWLAVGSMCLFIAGFVAGWGPVPLMSEIPLPHVKGVATGICVLTNWL 419
 Db 360 AAEPPVDVQVGLAWLAVGSMCLFIAGFVAGWGPVPLMSEIPLPHVKGVATGICVLTNWF 419
 QY 420 MAPLVTKEFSSLMVLRPYGAFWLASAFICFISVLFVLCVPEPKTKLQITAHFEGR 477
 Db 420 MAPLVTKEFNSVMEILRPYGAFWLTAFCILSVLFTLTPVPEPKTKLQITAHFEGR 477

RESULT 7
 ABM84862
 ID ABM84862 standard; protein; 408 AA.
 AC ABM84862;
 XX AC
 XX DT 18-NOV-2004 (first entry)
 XX DE Human diagnostic and therapeutic pprotein SEQ ID NO:5111.
 XX KW Gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
 XX OS Homo sapiens.
 XX PN WO2004023973-A2.
 XX PD 25-MAR-2004.
 XX PF 12-SEP-2003; 2003WO-US028227.
 XX PR 12-SEP-2002; 2002US-0410259P.
 XX PR 12-SEP-2002; 2002US-0410260P.
 XX PA (INCY-) INCYTE CORP.
 XX PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
 PI Hartschorne TA, Suchorski MT, Altus CM, Pitts SJ, Elder LV;
 PI Mooney EM, Deleage AM, Panesar IS, Banville SC, Reddy RP;
 PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;

QY 61 APRLDDAAASWFGAVVTLGAAAGVGLGWLDVDRAGKLSLLLSGVVFPVAGFAVITAAQDV 120
 Db 61 ALRLGDTAASWFGAVVTLGAAAGVGLGWLDVDRAGKLSLLLSGVVFPVAGFAVITAAQDV 120
 QY 121 WMLLGGRLLTGLACGVSALVAPVYIASEIAYPAVRGLLGSVCVQLMVTGILLAYVAGWVLE 180
 Db 121 WMLLGGRLLTGLACGVSALVAPVYIASEIAYPAVRGLLGSVCVQLMVTGILLAYVAGWVLE 180
 QY 181 WRWLAVLGCVPVPSLMLLMCFMPEPRFLLTQHRROEAMAAALRFLMGSEGGWEDPPIGAE 240
 Db 181 WRWLAVLGCVPVPSLMLLMCFMPEPRFLLTQHRROEAMAAALRFLMGSEGGWEDPPIGAE 240
 QY 241 -QSFLHALLRQPGIYKPFITIGVSLMAFOQLSGVNAVMFYAEITFEBAKFKDSSLASVVVG 299
 Db 241 HQGFQALMLRRPGVHKPLIIGICLMVYFQQLSGVNAIMFYANTIFEBAKFKDSSLASVVVG 299
 QY 300 VIQVLFATAALIMDRAGRRLLLVLSGVVMVFTSAFGAYFKLITQGGPNSSHVAISAPV 359
 Db 301 IIQVLFATAALIMDRAGRRLLLVLSGVVMVFTSAFGAYFKLITQGGPNSSHVAISAPV 359
 QY 360 SAQPVDAVGLAWLAVGSMCLFIAGFVAGWGPVPLMSEIPLPHVKGVATGICVLTNWL 419
 Db 361 SAEPADVHLGLAWLAVGSMCLFIAGFVAGWGPVPLMSEIPLPHVKGVATGICVLTNWF 420
 QY 420 MAPLVTKEFSSLMVLRPYGAFWLASAFICFISVLFVLCVPEPKTKLQITAHFEGR 477
 Db 421 MAPLVTKEFNSVMEILRPYGAFWLTAFCILSVLFTLTPVPEPKTKLQITAHFEGR 478

RESULT 6
 AAB66934
 ID AAB66934 standard; protein; 477 AA.
 AC AAB66934;
 XX AC
 XX DT 17-APR-2001 (first entry)
 XX DE Murine GLUTX1.
 XX KW Murine; GLUTX; gene therapy; vaccine; hexose transport modulator;
 KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia;
 KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 XX OS Mus sp.
 XX PN WO200104145-A2.
 XX PD 18-JAN-2001.
 XX PF 14-JUL-2000; 2000WO-IB001042.
 XX PR 14-JUL-1999; 99US-0143907P.
 XX PR 27-AUG-1999; 99US-0151140P.
 XX PR 13-FEB-2000; 2000US-0184285P.
 XX PR 13-JUL-2000; 2000US-00616132.
 XX PA (UYLA-) UNIV LAUSANNE.
 XX PI Thorens B, Ibberson M, Uldry M;
 XX DR WPI; 2001-112615/12.
 XX DR N-PSDB; AAF55867.
 XX PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in the
 PT prevention, diagnosis and treatment of hexose transport disorders, e.g.
 PT ischaemia and diabetes.
 XX PS Claim 14; Page 73-74; 124pp; English.
 XX CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
 CC AAB66932-AAB69411). The GLUTX proteins are related to the facultative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport

PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
 PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
 PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
 PI Patury S, Shi X, Suarez CJ;
 XX
 XX WPI: 2004-329368/30.
 DR N-PSDB; ACN43514.
 XX
 XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
 PT in diagnosing a condition, disease or disorder associated with human
 PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
 PT in gene mapping.
 XX
 XX Claim 27; Page: 190pp; English.
 XX
 XX The invention relates to novel diagnostic and therapeutic polynucleotides
 CC selected from one of the 2722 sequences defined in the specification. A
 CC polynucleotide of the invention may have a use in gene therapy. The human
 CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
 CC used to diagnose a particular condition, disease or disorder associated
 CC with human molecules, e.g. cell proliferative disorders,
 CC autoimmune/inflammatory disorder, developmental disorder, endocrine
 CC disorder, neurological disorders, gastrointestinal disorders, or
 CC infections caused by virus, bacteria, fungi or parasite. The dithp
 CC molecules may also be used in genetic mapping, in identifying individuals
 CC from minute biological samples, in detecting single nucleotide
 CC polymorphisms, as molecular weight markers, and for somatic or germline
 CC gene therapy. The present sequence represents a dithp protein of the
 CC invention. Note: The sequence data for this patent is not represented in
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
 XX
 XX Sequence 408 AA;
 DR Query Match 84.1%; Score 2067.5; DB 8; Length 408;
 DR Best Local Similarity 85.5%; Pred. No. 1.8e-189;
 DR Matches 408; Conservative 0; Mismatches 0; Indels 69; Gaps 1;

QY 1 MTPDEPEETQPLGPPGGSAPRRVFLAAFAAALGFLSFGALGYSSPAISLQRAAPP 60
 DB 1 MTPDEPEETQPLGPPGGSAPRRVFLAAFAAALGFLSFGALGYSSPAISLQRAAPP 60
 QY 61 APRLDDAASWFGAVVTLGAAAGVGLGWLDVDRGRKLSLLCCSVPFVAGFAVITAAQDV 120
 DB 61 APRLDDAASWFG-----VYISEIAYPAVRGLLGSCVQLMVVVGLLAYLAGWVLE 111
 QY 121 WMLLGGRLTGLACGVAASLVAIVYISEIAYPAVRGLLGSCVQLMVVVGLLAYLAGWVLE 180
 DB 74 -----VYISEIAYPAVRGLLGSCVQLMVVVGLLAYLAGWVLE 111
 QY 181 WRWLVLCVYPSIMLLMCMFETPRFLLTQHRROEAAALRFLWGSQGWEDPPIGAE 240
 DB 112 WRWLVLCVYPSIMLLMCMFETPRFLLTQHRROEAAALRFLWGSQGWEDPPIGAE 171
 QY 241 QSFHALLRPGYKPIIGVSLMAFQQLSGVNVAFYAEITFEAKFKDSSLASVYVGV 300
 DB 172 QSFHALLRPGYKPIIGVSLMAFQQLSGVNVAFYAEITFEAKFKDSSLASVYVGV 231
 QY 301 IQVLFATAALIMDRAGRLLLVLSGVVVFSTSAFGAYFKLTOGGFGNSHVAISAPYS 360
 DB 232 IQVLFATAALIMDRAGRLLLVLSGVVVFSTSAFGAYFKLTOGGFGNSHVAISAPYS 291
 QY 361 AQPVDASVGLAWLAVGSMCLFIAGFVGMGPIPWLLMSEIFPLHVKGVATGICVLINWLM 420
 DB 292 AQPVDASVGLAWLAVGSMCLFIAGFVGMGPIPWLLMSEIFPLHVKGVATGICVLINWLM 351
 QY 421 AFLVTKFSSIMEVLRVRYGAFWLASAFCIFSVLFTLFCVPEPKTKLEQITAHFEGR 477
 DB 352 AFLVTKFSSIMEVLRVRYGAFWLASAFCIFSVLFTLFCVPEPKTKLEQITAHFEGR 408

108 ID ARB84863 standard; protein; 382 AA.
 109 XX AEM84863;
 110 AC
 111 XX 18-NOV-2004 (first entry)
 112 DT
 113 DE Human diagnostic and therapeutic pprotein SEQ ID NO:5112.
 114 XX gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
 115 KW Homo sapiens.
 116 XX WO2004023973-A2.
 117 OS
 118 XX 25-MAR-2004.
 119 PD
 120 XX 12-SEP-2003; 2003WO-US028227.
 121 PF
 122 XX 12-SEP-2002; 2002US-0410259P.
 123 PR
 124 XX 12-SEP-2002; 2002US-0410260P.
 125 PR
 126 XX (INCY-) INCYTE CORP.
 127 PA
 128 XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F,
 129 XX Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
 130 XX Mooney EM, Deleane AM, Panesar IS, Banville SC, Reddy TP;
 131 XX Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
 132 XX Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
 133 XX Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
 134 XX Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
 135 XX Patury S, Shi X, Suarez CJ;
 136 XX
 137 DR WPI: 2004-329368/30.
 138 DR N-PSDB; ACN43515.
 139 XX
 140 XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
 141 PT in diagnosing a condition, disease or disorder associated with human
 142 PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
 143 PT in gene mapping.
 144 XX
 145 PS Claim 27; Page: 190pp; English.
 146 XX
 147 CC The invention relates to novel diagnostic and therapeutic polynucleotides
 148 CC selected from one of the 2722 sequences defined in the specification. A
 149 CC polynucleotide of the invention may have a use in gene therapy. The human
 150 CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
 151 CC used to diagnose a particular condition, disease or disorder associated
 152 CC with human molecules, e.g. cell proliferative disorders,
 153 CC autoimmune/inflammatory disorder, developmental disorder, endocrine
 154 CC disorder, neurological disorders, gastrointestinal disorders, or
 155 CC infections caused by virus, bacteria, fungi or parasite. The dithp
 156 CC molecules may also be used in genetic mapping, in identifying individuals
 157 CC from minute biological samples, in detecting single nucleotide
 158 CC polymorphisms, as molecular weight markers, and for somatic or germline
 159 CC gene therapy. The present sequence represents a dithp protein of the
 160 CC invention. Note: The sequence data for this patent is not represented in
 161 CC the printed specification, but was obtained in electronic format directly
 162 CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
 163 XX
 164 XX Sequence 382 AA;
 165 DR Query Match 77.8%; Score 1912.5; DB 8; Length 382;
 166 DR Best Local Similarity 80.1%; Pred. No. 1.2e-174;
 167 DR Matches 382; Conservative 0; Mismatches 0; Indels 95; Gaps 2;

QY 1 MTPDEPEETQPLGPPGGSAPRRVFLAAFAAALGFLSFGALGYSSPAISLQRAAPP 60
 DB 1 MTPDEPEETQPLGPPGGSAPRRVFLAAFAAALGFLSFGALGYSSPAISLQRAAPP 60
 QY 61 APRLDDAASWFGAVVTLGAAAGVGLGWLDVDRGRKLSLLCCSVPFVAGFAVITAAQDV 120
 DB 61 APRLDDAASWFG-----VYISEIAYPAVRGLLGSCVQLMVVVGLLAYLAGWVLE 111
 QY 121 WMLLGGRLTGLACGVAASLVAIVYISEIAYPAVRGLLGSCVQLMVVVGLLAYLAGWVLE 180
 DB 74 -----VYISEIAYPAVRGLLGSCVQLMVVVGLLAYLAGWVLE 111
 QY 181 WRWLVLCVYPSIMLLMCMFETPRFLLTQHRROEAAALRFLWGSQGWEDPPIGAE 240
 DB 112 WRWLVLCVYPSIMLLMCMFETPRFLLTQHRROEAAALRFLWGSQGWEDPPIGAE 171
 QY 241 QSFHALLRPGYKPIIGVSLMAFQQLSGVNVAFYAEITFEAKFKDSSLASVYVGV 300
 DB 172 QSFHALLRPGYKPIIGVSLMAFQQLSGVNVAFYAEITFEAKFKDSSLASVYVGV 231
 QY 301 IQVLFATAALIMDRAGRLLLVLSGVVVFSTSAFGAYFKLTOGGFGNSHVAISAPYS 360
 DB 232 IQVLFATAALIMDRAGRLLLVLSGVVVFSTSAFGAYFKLTOGGFGNSHVAISAPYS 291
 QY 361 AQPVDASVGLAWLAVGSMCLFIAGFVGMGPIPWLLMSEIFPLHVKGVATGICVLINWLM 420
 DB 292 AQPVDASVGLAWLAVGSMCLFIAGFVGMGPIPWLLMSEIFPLHVKGVATGICVLINWLM 351
 QY 421 AFLVTKFSSIMEVLRVRYGAFWLASAFCIFSVLFTLFCVPEPKTKLEQITAHFEGR 477
 DB 352 AFLVTKFSSIMEVLRVRYGAFWLASAFCIFSVLFTLFCVPEPKTKLEQITAHFEGR 408

108 ID ARB84863 standard; protein; 382 AA.
 109 XX AEM84863;
 110 AC
 111 XX 18-NOV-2004 (first entry)
 112 DT
 113 DE Human diagnostic and therapeutic pprotein SEQ ID NO:5112.
 114 XX gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
 115 KW Homo sapiens.
 116 XX WO2004023973-A2.
 117 OS
 118 XX 25-MAR-2004.
 119 PD
 120 XX 12-SEP-2003; 2003WO-US028227.
 121 PF
 122 XX 12-SEP-2002; 2002US-0410259P.
 123 PR
 124 XX 12-SEP-2002; 2002US-0410260P.
 125 PR
 126 XX (INCY-) INCYTE CORP.
 127 PA
 128 XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F,
 129 XX Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
 130 XX Mooney EM, Deleane AM, Panesar IS, Banville SC, Reddy TP;
 131 XX Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
 132 XX Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
 133 XX Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
 134 XX Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
 135 XX Patury S, Shi X, Suarez CJ;
 136 XX
 137 DR WPI: 2004-329368/30.
 138 DR N-PSDB; ACN43515.
 139 XX
 140 XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
 141 PT in diagnosing a condition, disease or disorder associated with human
 142 PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
 143 PT in gene mapping.
 144 XX
 145 PS Claim 27; Page: 190pp; English.
 146 XX
 147 CC The invention relates to novel diagnostic and therapeutic polynucleotides
 148 CC selected from one of the 2722 sequences defined in the specification. A
 149 CC polynucleotide of the invention may have a use in gene therapy. The human
 150 CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
 151 CC used to diagnose a particular condition, disease or disorder associated
 152 CC with human molecules, e.g. cell proliferative disorders,
 153 CC autoimmune/inflammatory disorder, developmental disorder, endocrine
 154 CC disorder, neurological disorders, gastrointestinal disorders, or
 155 CC infections caused by virus, bacteria, fungi or parasite. The dithp
 156 CC molecules may also be used in genetic mapping, in identifying individuals
 157 CC from minute biological samples, in detecting single nucleotide
 158 CC polymorphisms, as molecular weight markers, and for somatic or germline
 159 CC gene therapy. The present sequence represents a dithp protein of the
 160 CC invention. Note: The sequence data for this patent is not represented in
 161 CC the printed specification, but was obtained in electronic format directly
 162 CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
 163 XX
 164 XX Sequence 382 AA;
 165 DR Query Match 77.8%; Score 1912.5; DB 8; Length 382;
 166 DR Best Local Similarity 80.1%; Pred. No. 1.2e-174;
 167 DR Matches 382; Conservative 0; Mismatches 0; Indels 95; Gaps 2;

CC signaling molecule, a naturally occurring amino acid sequence at least 90% identical to it or a biologically active fragment of an immunogenic fragment of the polypeptide. The human TRICH, polynucleotides, agonists and antagonists are useful for diagnosing, treating or preventing disorders associated with aberrant expression of TRICH, particularly cell proliferative disorders (e.g. arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocytopenia or cancer), developmental disorders (e.g. renal tubular acidosis, anemia or mental retardation), neurological disorders (e.g. Alzheimer's disease, Parkinson's disease or epilepsy), autoimmune/inflammatory disorders (e.g. AIDS, allergies, asthma, autoimmune thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, irritable bowel syndrome, multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis, Reiter's syndrome, rheumatoid arthritis, Sjogren's syndrome, uveitis), or viral, bacterial, fungal, parasitic, protozoan or helminthic infections. The polynucleotides encoding TRICH are useful for creating transgenic animals to model human disease. This sequence corresponds to one of the proteins of the invention.

XX SQ Sequence 353 AA;

Query Match 67.7%; Score 1664; DB 7; Length 353;
 Best Local Similarity 100.0%; Pred. No. 8e-151;
 Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTPDEPETQPLGPPGSSAPRGRVFLAAFAAALGPLSFGALGVSSPAISLQRAAPP 60
 DB 1 MTPDEPETQPLGPPGSSAPRGRVFLAAFAAALGPLSFGALGVSSPAISLQRAAPP 60

QY 61 APRLDDAASWEGAVVTLGAAAGVGLGGWLVDRAGKLSLLCSVFPFVAGFAVITAAQDV 120
 DB 61 APRLDDAASWEGAVVTLGAAAGVGLGGWLVDRAGKLSLLCSVFPFVAGFAVITAAQDV 120

QY 121 WMLLGGRLLTGLACGVASIVAPVYISETAYPAVRGLLSCVQLMVMVVGILLAYLAGWVLE 180
 DB 121 WMLLGGRLLTGLACGVASIVAPVYISETAYPAVRGLLSCVQLMVMVVGILLAYLAGWVLE 180

QY 181 WRWLAVIGCVPPSLMMLLMLCMEPEPRFLLTQHRQEAARLFLWGSQGWEDDPDGA 240
 DB 181 WRWLAVIGCVPPSLMMLLMLCMEPEPRFLLTQHRQEAARLFLWGSQGWEDDPDGA 240

QY 241 QSFHLALLRQPGIYKPFIIIGVSLMAFQOQLSGVNAVWFYAETIFEEAKFKDSSLASVVVGV 300
 DB 241 QSFHLALLRQPGIYKPFIIIGVSLMAFQOQLSGVNAVWFYAETIFEEAKFKDSSLASVVVGV 300

QY 301 IQVLFATAALIMDRAGRRLLLVLSG 326
 DB 301 IQVLFATAALIMDRAGRRLLLVLSG 326

CC signaling molecule, a naturally occurring amino acid sequence at least 90% identical to it or a biologically active fragment of an immunogenic fragment of the polypeptide. The human TRICH, polynucleotides, agonists and antagonists are useful for diagnosing, treating or preventing disorders associated with aberrant expression of TRICH, particularly cell proliferative disorders (e.g. arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocytopenia or cancer), developmental disorders (e.g. renal tubular acidosis, anemia or mental retardation), neurological disorders (e.g. Alzheimer's disease, Parkinson's disease or epilepsy), autoimmune/inflammatory disorders (e.g. AIDS, allergies, asthma, autoimmune thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, irritable bowel syndrome, multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis, Reiter's syndrome, rheumatoid arthritis, Sjogren's syndrome, uveitis), or viral, bacterial, fungal, parasitic, protozoan or helminthic infections. The polynucleotides encoding TRICH are useful for creating transgenic animals to model human disease. This sequence corresponds to one of the proteins of the invention.

XX SQ Sequence 353 AA;

Query Match 67.7%; Score 1664; DB 7; Length 353;
 Best Local Similarity 100.0%; Pred. No. 8e-151;
 Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTPDEPETQPLGPPGSSAPRGRVFLAAFAAALGPLSFGALGVSSPAISLQRAAPP 60
 DB 1 MTPDEPETQPLGPPGSSAPRGRVFLAAFAAALGPLSFGALGVSSPAISLQRAAPP 60

QY 61 APRLDDAASWEGAVVTLGAAAGVGLGGWLVDRAGKLSLLCSVFPFVAGFAVITAAQDV 120
 DB 61 APRLDDAASWEGAVVTLGAAAGVGLGGWLVDRAGKLSLLCSVFPFVAGFAVITAAQDV 120

QY 121 WMLLGGRLLTGLACGVASIVAPVYISETAYPAVRGLLSCVQLMVMVVGILLAYLAGWVLE 180
 DB 121 WMLLGGRLLTGLACGVASIVAPVYISETAYPAVRGLLSCVQLMVMVVGILLAYLAGWVLE 180

QY 181 WRWLAVIGCVPPSLMMLLMLCMEPEPRFLLTQHRQEAARLFLWGSQGWEDDPDGA 240
 DB 181 WRWLAVIGCVPPSLMMLLMLCMEPEPRFLLTQHRQEAARLFLWGSQGWEDDPDGA 240

QY 241 QSFHLALLRQPGIYKPFIIIGVSLMAFQOQLSGVNAVWFYAETIFEEAKFKDSSLASVVVGV 300
 DB 241 QSFHLALLRQPGIYKPFIIIGVSLMAFQOQLSGVNAVWFYAETIFEEAKFKDSSLASVVVGV 300

QY 301 IQVLFATAALIMDRAGRRLLLVLSG 326
 DB 301 IQVLFATAALIMDRAGRRLLLVLSG 326

RESULT 10
 ABB89717
 ID ABB89717 standard; protein; 326 AA.
 AC ABB89717;
 XX ABB89717;
 DT 24-MAY-2002 (first entry)
 XX Human polypeptide SEQ ID NO 2093.
 XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
 KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 OS Homo sapiens.
 XX WO200190304-A2.
 XX 29-NOV-2001.

RESULT 9
 ADL33342
 ID ADL33342 standard; protein; 353 AA.
 AC ADL33342;
 XX ADL33342;
 DT 20-MAY-2004 (first entry)
 XX Human transporter and ion channel (TRICH) protein #46.
 KW anti-HIV; anti-allergic; anti-inflammatory; antianemic; antiparkinsonian;
 KW nootropic; anticonvulsant; antiarteriosclerotic; antiaesthetic;
 KW immunosuppressive; antithyroid; cytostatic; hepatotropic; dermatological;
 KW antidiabetic; aephrotropic; antigout; thyromimetic; neuroprotective;
 KW osteopathic; antiarthritic; antiparasitic; antihelminthic; antiparasitic;
 KW uropathic; ophthalmological; antirheumatic; hemostatic; antibacterial;
 KW virucide; protozoacide; fungicide; gene therapy.
 XX Homo sapiens.
 OS WO2003083085-A2.
 XX 09-OCT-2003.
 XX 27-MAR-2003; 2003WO-US0009797.
 XX 28-MAR-2002; 2002US-0368840P.
 XX 26-APR-2002; 2002US-0375637P.
 XX (INCY-) INCYTE CORP.
 XX Marquis JP, Lee SY, Emerling BM, Hafalia AJA, Khare R, Kable AB;
 PI Richardson TW, Swarnakar A, Chawla NK, Becha SD, Mason PM;
 PI Elliott VS, Ramkumar J, Griffin JA, Tran UK, Ison CH, Lindquist EA;
 PI Jiang X, Jackson AA, Wilson AD, Jin P, Chang H;
 XX WPI; 2003-833535/77.
 DR N-PSDB; ADL33401.
 XX New human transporters and ion channels (TRICH) and polynucleotides,
 PT useful for diagnosing, treating or preventing autoimmune or inflammatory
 PT disorders (e.g. AIDS, allergy or anemia), multiple sclerosis, cancer or
 PT hepatitis.
 XX Claim 1; SEQ ID NO 46; 405pp; English.
 PS The invention relates to an isolated polypeptide (I), which is a human
 CC intracellular signaling molecule, which is a human intracellular

XX 18-MAY-2001; 2001WO-US016450.
 XX 19-MAY-2000; 2000US-0205515P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Birse CB, Rosen CA;
 XX WPI; 2002-122018/16.
 XX N-PSDB; ABL90126.
 XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 XX prevention of neural, immune system, muscular, reproductive,
 XX gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 XX disorders.
 XX Claim 11; SEQ ID NO 2093; 2081pp + Sequence Listing; English.
 XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
 XX (ABB9040-ABB9044) useful for preventing, treating or ameliorating
 XX medical conditions e.g. by protein or gene therapy. The genes are
 XX isolated from a range of human tissues disclosed in the specification.
 XX The nucleic acids, proteins, antibodies and (ant)agonists are useful in
 XX the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 XX ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
 XX breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
 XX disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 XX cardiovascular disorders such as myocardial ischaemias; (d) wound healing
 XX ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 XX infectious diseases such as viral, bacterial, fungal and parasitic
 XX infections. Note: The sequence data for this patent did not form part of
 XX the printed specification, but was obtained in electronic format directly
 XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 326 AA;

Query Match 60.7%; Score 1491; DB 5; Length 326;
 Best Local Similarity 96.3%; Pred. No. 2.9e-134;
 Matches 289; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 164 MVVVGILLAYLAGWLEWELAVLGCVPSSLMLLMLCFMPEPRFLLTQHRREQEWAALR 223
 1 MVVVGILLAYLAGWLEWELAVLGCVPSSLMLLMLCFMPEPRFLLTQHRREQEWAALR 60

QY 224 FLWGSQGWEDPPIGAEQSFHLALRQPGIYKPFIIIGVSLMAFQQLSGVNAVWYAEITF 283
 61 FLWGSQGWEDPPIGAEQSFHLALRQPGIYKPFIIIGVSLMAFQQLSGVNAVWYAEITF 120

QY 284 EAKFKDSSLASVVVGVIOVLFVAALIMDRAGRLLLVLSGVVWVFTSAFGYFKLT 343
 121 EAKFKDSSLASVVVGVIOVLFVAALIMDRAGRLLLVLSGVVWVFTSAFGYFKLT 180

QY 344 QGPGNSSHVAISAPYSAGPVDASVGLAVGSMCLFTAGFAGVGPPIWLLMSEIRPL 403
 181 QGPGNSSHVAISAPYSAGPVDASVGLAVGSMCLFTAGFAGVGPPIWLLMSEIRPL 240

QY 404 HVKGVATGICVLNTNLMALVTKKFSLSMEVLRPYGAPFWLASAFCLFSLFTLFCVPEPK 463
 241 HVKGVATGICVLNTNLMALVTKKFSLSMEVLRPYGAPFWLASAFCLFSLFTLFCVPEPK 300

RESULT 11
 AAE06579
 ID AAE06579 standard; protein; 262 AA.
 XX AAE06579;
 XX 25-SEP-2001 (first entry)
 XX Human protein having hydrophobic domain, HP10784.

XX Human; hydrophobic domain; gene therapy; nutritional supplement;
 XX cell proliferation; immunomodulatory; autoimmune disorder; antimicrobial;
 XX multiple sclerosis; rheumatoid arthritis; insulin-dependent diabetes;
 XX haematopoiesis; tissue growth activity; Parkinson's disease; cytostatic;
 XX Huntington's disease; Alzheimer's disease; chemotactic; chemokinetic;
 XX haemostatic; thrombolytic; tumour growth inhibitor; anabolic;
 XX contraceptive; antiinfertility; antiinflammatory.
 XX Homo sapiens.
 XX WO200149728-A2.
 XX 12-JUL-2001.
 XX 28-DEC-2000; 2000WO-JP009359.
 XX 06-JAN-2000; 2000JP-0000585.
 XX 06-JAN-2000; 2000JP-0000588.
 XX 11-JAN-2000; 2000JP-00002299.
 XX 03-FEB-2000; 2000JP-00026862.
 XX 03-MAR-2000; 2000JP-00058367.
 XX (PROT-) PROTEGENE INC.
 XX (SAGA) SAGAMI_CHEM RES CENT.
 XX Kato S, Kimura T;
 XX WPI; 2001-418355/44.
 XX N-PSDB; AAD12574.
 XX Human proteins with hydrophobic domains and the nucleic acids encoding
 XX them, useful for preventing diagnosing and treating e.g. cancer,
 XX Alzheimer's and inflammation.
 XX Claim 1; Page 75; 563pp; English.

XX The present sequence is human protein with hydrophobic domain, HP10784.
 XX The polynucleotide and polypeptide of the invention may be used in the
 XX prevention, diagnosis and treatment of diseases associated with
 XX inappropriate polypeptide expression. The polynucleotides may be used to
 XX produce the polypeptide, by inserting the nucleic acids into a host cell
 XX and culturing the cell to express the protein. The polynucleotides and
 XX its complementary sequences may also be used as DNA probes in diagnostic
 XX assays and also used in gene therapy. The polypeptides may also be used
 XX as antigens in the production of antibodies and in assays to identify
 XX modulators of polypeptide expression and activity. The polypeptides and
 XX nucleic acids may be used as nutritional supplements, to modulate
 XX cytokine and cell proliferation activity, to modulate immune stimulation
 XX or suppression (e.g. for the treatment of microbial infections and
 XX autoimmune disorders such as multiple sclerosis, rheumatoid arthritis and
 XX insulin-dependent diabetes), to modulate haematopoiesis to modulate
 XX tissue growth activity (e.g. for the treatment of Parkinson's disease,
 XX Huntington's disease and Alzheimer's disease), to modulate activin and
 XX inhibin activity (e.g. for controlling fertility), to modulate
 XX chemotactic and chemokinetic activity, to modulate haemostatic and
 XX thrombolytic activity, to modulate receptor ligand activity, to modulate
 XX inflammation and to inhibit tumour growth

XX SQ Sequence 262 AA;
 Query Match 46.2%; Score 1135; DB 4; Length 262;
 Best Local Similarity 91.4%; Pred. No. 3.3e-100;
 Matches 223; Conservative 2; Mismatches 7; Indels 12; Gaps 1;

QY 1 MTPEDPETQLLGGPPGSGAPRRVFLAAFAAALPLSFGFALGYSSPATSLSORAPP 60
 1 MTPEDPETQLLGGPPGSGAPRRVFLAAFAAALPLSFGFALGYSSPATSLSORAPP 60

QY 61 APRLDDAAASWFGAVVTLGAAAGGVLGGWLVDRAGKLSLLLCVFPVFAVITAAQDV 120
 61 APRLDDAAASWFGAVVTLGAAAGGVLGGWLVDRAGKLSLLLCVFPVFAVITAAQDV 120

CC for GLUTX3
 XX Sequence 507 AA;
 SQ Query Match 38.8%; Score 953; DB 4; Length 507;
 Best Local Similarity 42.9%; Pred. No. 2.4e-82;
 Matches 217; Conservative 78; Mismatches 167; Indels 44; Gaps 11;

QY 10 QPLLG-----PFGSAPRGR-----RVLEAAPAAALGPLSFGFALGYSSPA 50
 Db 3 EPLLGAEQPDYTFPEKPPASPGDRARVGTQNKRVFLATFAVLGNFSGFALVYTFPV 62
 QY 51 IPGLQRAAPPAPRLDDAAASWFGAVVTLGAAAGGVLGWLVDRAQRKLSLLLCSPVFFVAG 110
 Db 63 IPALERSLPDLHLTKSQASWFGSVFTLGAAGGLSAMTLNDLLGRKLSIMFSAPVSAAG 122
 QY 111 FAVITAAQDVMMLLGGRLTLGLACGVASLVAPYIYSEIAYPAVRGLLSCVQVMVVVGL 170
 Db 123 YAVMAGAHGLMMLLGRITLTGFAGGLTAAACIFVYVSEIAPPVGRGALGATPQLMAVFGSL 182
 QY 171 LAYLWVLEWVLAIVLGCVPSPSLMLLMLLMLLMLLMLLMLLMLLMLLMLLMLLMLL 230
 Db 183 SLYALGILLPWRWLVAVAGKAPVLMILLMLLMLLMLLMLLMLLMLLMLLMLLMLLMLL 242
 QY 231 G--WEDPPPIG--AEQS FHL--ALLRQPGIYKPFIIIGVSLMAFPQQLSGVNVMFYAEITF 283
 Db 243 DVHWEFGQIQDNVRRQSSRVSWAERAPHVCRPITVALLMLLMLLMLLMLLMLLMLL 302
 QY 284 EEAKF----KSSLASVWVGVIVFTVAALIMDRAGRLMLLMLLMLLMLLMLLMLLMLL 339
 Db 303 DSTAVLLPPKDD---AAIVGAVRLLSVLTAALTMLAGRKLVLVFSAAALMFAANLTLGLY 359
 QY 340 FKLTOGGPGNSHVAISAPVS-----AOPVDASVG-LAWLAVGSMCLFIAGFVAGWGPI 392
 Db 360 IHF---GPRRLSPNSTAGLESWGDLAQPLAAPAGVLTIVPLLATMLTFLMGYAVGWGPI 416
 QY 393 PWLMSIFPLHVKVATGICVTLWMLMFLVTKFESSLMEVLRFPYGAFLASAFICFSV 452
 Db 417 TWLIMSEVPLRARGVNASGLCVLASMWLTAFVLTKSEFLPVVSTFGLQVPPFLFAAICVSL 476
 QY 453 LFTLFCVPETKGTLEQITAHFB-GR 477
 Db 477 VFTGCCVPETKGRSLEQIESFFRTGR 502

RESULT 15
 AAB66937
 ID AAB66937 standard; protein; 507 AA.
 XX
 AC AAB66937;
 XX
 DT 17-APR-2001 (first entry)
 XX
 DE Human GLUTX3.
 XX
 KW Human; GLUTX; gene therapy; vaccine; hexose transport modulator;
 KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia;
 KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 XX
 OS Homo sapiens.
 XX
 XX WO200104145-A2.
 XX
 XX 18-JAN-2001.
 XX
 XX 14-JUL-2000; 2000WO-IB001042.
 XX
 XX 14-JUL-1999; 99US-0143907P.
 XX 27-AUG-1999; 99US-0151140P.
 XX 23-FEB-2000; 2000US-0184285P.
 XX 13-JUL-2000; 2000US-00616132.
 XX
 XX (UYLA-) UNIV LAUSANNE.
 XX
 XX Thorens B, Ibberson M, Uldry M;
 XX
 XX WPI; 2001-112615/12.
 XX
 XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in the
 XX prevention, diagnosis and treatment of hexose transport disorders, e.g.
 XX ischemia and diabetes.
 XX
 XX Claim 11; Page 83-84; 124pp; English.
 XX
 XX The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
 XX AAB66932-AAB66941). The GLUTX proteins are related to the facultative
 XX glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 XX function. The GLUTX proteins may be used in the diagnosis, prevention and
 XX treatment of hexose transport disorders such as ischaemia, diabetes,
 XX hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 XX neurodegenerative disease. The present sequence is a consensus sequence

QY 183 WLAVLGCVPSPSLMLLMLLMLLMLLMLLMLLMLLMLLMLLMLLMLLMLLMLLMLLMLL 238
 Db 194 WLAVAGEFVLMVILLLSFMPNSPRFLLSKSRDEALQAL--IWLADSEVHWFEEQIQD 251
 QY 239 --AEQSFHALLR--OPGIYKPFIIIGVSLMAFPQQLSGVNVMFYAEITFBEAK-FKDS 293
 Db 252 NVRRQSSRVSWAERAPHVCRPITVALLMLLMLLMLLMLLMLLMLLMLLMLLMLLMLL 311
 QY 294 ASVVVGVIVFTVAALIMDRAGRLMLLMLLMLLMLLMLLMLLMLLMLLMLLMLLMLL 351
 Db 312 DAAIVGAVRLLSVLTAALTMLAGRKLVLVFSAAALMFAANLTLGLYVQVPRTLTNSIV 371
 QY 352 HVAISAPVSAQVDSVGLAWLAVGSMCLFIAGFVAGWGPIPWLMSIFPLHVKVATG 411
 Db 372 EIVTLGGTEPAAAFNYLTLIPLLATMLTFLMGYAMGWGITWILMSEVPLRARGVAS 431
 QY 412 ICVLTWMLMFLVTKFESSLMEVLRPYGAFWLASAFICFVSLFTLFCVPETKGTLEQIT 471
 Db 432 LCVLVSWLTAFAVLTXYLLAVNAFLQVPPFFSALCLLSLLFTGCCVPETKGRSLEQIE 491
 QY 472 AHFEGR 477
 Db 492 ARFHTR 497

RESULT 14
 AAB66941
 ID AAB66941 standard; protein; 507 AA.
 XX
 AC AAB66941;
 XX
 DT 17-APR-2001 (first entry)
 XX
 DE GLUTX3 consensus sequence.
 XX
 KW GLUTX; gene therapy; vaccine; hexose transport modulator; human; rat;
 KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia;
 KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 XX
 OS Homo sapiens.
 XX
 XX Rattus sp.
 XX
 XX WO200104145-A2.
 XX
 XX 18-JAN-2001.
 XX
 XX 14-JUL-2000; 2000WO-IB001042.
 XX
 XX 14-JUL-1999; 99US-0143907P.
 XX 27-AUG-1999; 99US-0151140P.
 XX 23-FEB-2000; 2000US-0184285P.
 XX 13-JUL-2000; 2000US-00616132.
 XX
 XX (UYLA-) UNIV LAUSANNE.
 XX
 XX Thorens B, Ibberson M, Uldry M;
 XX
 XX WPI; 2001-112615/12.
 XX
 XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in the
 XX prevention, diagnosis and treatment of hexose transport disorders, e.g.
 XX ischemia and diabetes.
 XX
 XX Claim 11; Page 83-84; 124pp; English.
 XX
 XX The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
 XX AAB66932-AAB66941). The GLUTX proteins are related to the facultative
 XX glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 XX function. The GLUTX proteins may be used in the diagnosis, prevention and
 XX treatment of hexose transport disorders such as ischaemia, diabetes,
 XX hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 XX neurodegenerative disease. The present sequence is a consensus sequence

XX Thorens B, Ibberson M, Uldry M;
 XX WPI; 2001-112615/12.
 DR N-PSDB; AAF55870.
 XX
 PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in the
 PT prevention, diagnosis and treatment of hexose transport disorders, e.g.
 PT ischemia and diabetes.
 XX
 PS Claim 11; Page 81-82; 124pp; English.
 XX
 CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facultative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is human GLUTX3
 XX
 SQ Sequence 507 AA;
 Query Match ; 38.6%; Score 948; DB 4; Length 507;
 Best Local Similarity 42.9%; Pred. No. 7.1e-82;
 Matches 217; Conservative 78; Mismatches 167; Indels 44; Gaps 11;

QY 10 QPLLG-----PFGGSAPRGR-----RVFLAAFAAALGLPFGALGYSSPA 50
 Db 3 EPLLAGEGPDYDFPEKPPSPGDRARVGTLQNKRVFLATFAAVLGNFSGYALVYTSVP 62

QY 51 IFSLQRAAPPAPRLDDAASWFGAVVTLGAAAGVLLGGWLVDRAGRKLSELLCSVPEFVAG 110
 Db 63 IPALERSLDPDLHLTKSQASWFGSVFTLGAAGLSAMTINDLLGRKLSIMFSAVPSAAG 122

QY 111 FAVITRAQDVMLLGGRLTGLACGVASLVAVPVYISEIYPAVRGLLSCVOLMNVVVGIL 170
 Db 123 YAVMAGAHGLMILLGRTLTGFAGLTAACIPVYVSEIAPPVGRGALGATPQLMAVFGSL 182

QY 171 LAYLAGWLEWRWLAVLGCVPPLMLLLMCFPEPRFLTLTOHRRQEAALRFLMGSEQ 230
 Db 183 SLVALGULLPWRWLAVAGXAPVLIIMILLISFMPNSPFLLSRGRDEALRALAWLREGTDV 242

QY 231 G--WEDPPIG--AEQSFHL--ALLRQPGHYKPFIIIGVSLMAFQQLSGVNAVNFYABTIF 283
 Db 243 DVHWERFQIQDNVRRQSRVSWAEARAPHVCRPITVALLMRLQLQLTGITPILVYLQSI 302

QY 284 BEAKP---KSSLASVVVGVIOVFTVAALIMDRGRLLLVLSGVVWFSTSAFGAY 339
 Db 303 DSTAVLLPKDD---AAIVGXVRLLSVLIATLMDLAGRKKVLLFVSAAIMFAANLTLGLY 359

QY 340 FKLTOGGPGNSHVAISAPVS-----AQPVDASVG-LAWLAVGSMCLFIAGFVGVGPI 392
 Db 360 IHF---GPRPLSPNSTAGLESWGDLAQLAPAGYLTIVPLLATMFLIMGYAVGVGPI 416

QY 393 PMLMSEIIFPLHVKGVATGICVLTNWMAFLVTKFESSLMVEVLRPYGAFWLASAFCFISV 452
 Db 417 TWLLMSEVLPARGVAVGLCVLASMLTAFVLTSLFVVSSTFGLQVPLFFAAICLVSL 476

QY 453 LFTLFCVPEPKTLEQITAFEB-GR 477
 Db 477 VFTGCCVPETKGRSLQIESFFRTGR 502

Search completed: February 24, 2005, 05:49:51
 Job time : 175 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 05:47:06 ; Search time 132 Seconds
(without alignments)
1182.530 Million cell updates/sec

Title: US-09-886-954A-1
Perfect score: 2457
Sequence: 1 MTPPEPEFQPLGPPGSA.....CVPETKGTLEQITAHFGR 477

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1380268 seqs, 327241040 residues 1380268

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Table with columns: Result No., Score, Match, Length, DB ID, Description. Contains 20 rows of search results.

Table with columns: 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45. Contains alignment data.

ALIGNMENTS

RESULT 1
US-10-168-651-1
Sequence 1, Application US/10168651
Publication No. US20030171275A1
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: BAUGHN, Mariah R.
APPLICANT: BURFORD, Neil
APPLICANT: AU-YOUNG, Janice
APPLICANT: LU, Dyung Aina M.
APPLICANT: YANG, Junming
APPLICANT: REDDY, Roopa
APPLICANT: LAL, Preeti
APPLICANT: HILLMAN, Jennifer L.
APPLICANT: AZIMZAI, Yalda
APPLICANT: YUE, Henry
APPLICANT: NGUYEN, Dannie B.
APPLICANT: YAO, Monique G.
APPLICANT: GANDHI, Ameeni R.
APPLICANT: TANG, Y. Tom
APPLICANT: KHAN, Farrah A.
TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS
FILE REFERENCE: PI-0005 PCT
CURRENT APPLICATION NUMBER: US/10/168,651
PRIOR FILING DATE: 2002-06-21
PRIOR APPLICATION NUMBER: 60/172,000; 60/176,083; 60/177,332; 60/178,572; 60/179,758
60/181,625
PRIOR FILING DATE: 1999-12-23; 2000-01-14; 2000-01-21; 2000-01-28; 2000-02-02; 2000-02-10
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PERL Program
SEQ ID NO 1
LENGTH: 477
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:

NAME/KEY: misc feature
 OTHER INFORMATION: Incyte ID No. US20030171275A1 1416107CD1
 US-10-168-651-1

Query Match 100.0%; Score 2457; DB 14; Length 477;
 Best Local Similarity 100.0%; Pred. No. 2.9e-203;
 Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTPEDPEETQPLGGPPGAPRRVFLAAFAAALGFLSFGFALGYSSPAIPSLQRAAPP 60
 DB 1 MTPEDPEETQPLGGPPGAPRRVFLAAFAAALGFLSFGFALGYSSPAIPSLQRAAPP 60

QY 61 APRLDDAAASWFGAVVTLGAAAGVGLGGWLVDRAGKLSLLCSVFFVAGFAVITAAQDV 120
 DB 61 APRLDDAAASWFGAVVTLGAAAGVGLGGWLVDRAGKLSLLCSVFFVAGFAVITAAQDV 120

QY 121 WMLLGRLLTGLACGVASLVAFYISETAYPAVRGLLGCVCVQVMMVVVGLLAYLAGWVLE 180
 DB 121 WMLLGRLLTGLACGVASLVAFYISETAYPAVRGLLGCVCVQVMMVVVGLLAYLAGWVLE 180

QY 181 WRWLAVLGCVPVPSLMLLMCFMPEPRFLLTQHRROEAMAAALRFLWGSEQGWEDDPPIGAE 240
 DB 181 WRWLAVLGCVPVPSLMLLMCFMPEPRFLLTQHRROEAMAAALRFLWGSEQGWEDDPPIGAE 240

QY 241 QSHLALLRQPGIYKPFIIIGVSIIMAFQQLSGVNAVNFYAEITIFEAKFKDSSLASVWVGV 300
 DB 241 QSHLALLRQPGIYKPFIIIGVSIIMAFQQLSGVNAVNFYAEITIFEAKFKDSSLASVWVGV 300

QY 301 IOVLFTVAALIMDRAGRLLLVLSGVVWVFTSFAFGYFKLTQGGPGNSHVAISAPVS 360
 DB 301 IOVLFTVAALIMDRAGRLLLVLSGVVWVFTSFAFGYFKLTQGGPGNSHVAISAPVS 360

QY 361 AQPVDASVGLAWLAVGSMCLFIAGFAVWGPIPWLLMSEIFPLHVKGVAATGICVLTNMLM 420
 DB 361 AQPVDASVGLAWLAVGSMCLFIAGFAVWGPIPWLLMSEIFPLHVKGVAATGICVLTNMLM 420

QY 421 AFLVTKFSSLMVLRPYGAFWLASAFCSFVSLFTLFCVPEYKTKLEQITAHFEGR 477
 DB 421 AFLVTKFSSLMVLRPYGAFWLASAFCSFVSLFTLFCVPEYKTKLEQITAHFEGR 477

RESULT 3

US-10-264-237-2093
 Sequence 2093, Application US/10264237
 Publication No. US20040009491A1
 GENERAL INFORMATION:
 APPLICANT: Btise et al.
 TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 FILE REFERENCE: PA131P1
 CURRENT APPLICATION NUMBER: US/10/264,237
 CURRENT FILING DATE: 2002-10-04
 PRIOR APPLICATION NUMBER: PCT/US01/16450
 PRIOR FILING DATE: 2001-05-18
 PRIOR APPLICATION NUMBER: US 60/205,515
 PRIOR FILING DATE: 2000-05-19
 NUMBER OF SEQ ID NOS: 2876
 SOFTWARE: Patent In Ver. 3.1
 SEQ ID NO 2093
 LENGTH: 326
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: MISC_FEATURE
 LOCATION: (76)
 OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
 FEATURE:
 NAME/KEY: MISC_FEATURE
 LOCATION: (87)
 OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
 FEATURE:
 NAME/KEY: MISC_FEATURE
 LOCATION: (249)
 OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
 FEATURE:
 NAME/KEY: MISC_FEATURE
 LOCATION: (253)
 OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
 FEATURE:
 NAME/KEY: MISC_FEATURE
 LOCATION: (258)
 OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
 FEATURE:
 NAME/KEY: MISC_FEATURE
 LOCATION: (261)
 OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
 FEATURE:
 NAME/KEY: MISC_FEATURE
 LOCATION: (316)
 OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
 US-10-264-237-2093

NAME/KEY: misc feature
 OTHER INFORMATION: Incyte ID No. US20030171275A1 1416107CD1
 US-10-168-651-1

Query Match 100.0%; Score 2457; DB 14; Length 477;
 Best Local Similarity 100.0%; Pred. No. 2.9e-203;
 Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTPEDPEETQPLGGPPGAPRRVFLAAFAAALGFLSFGFALGYSSPAIPSLQRAAPP 60
 DB 1 MTPEDPEETQPLGGPPGAPRRVFLAAFAAALGFLSFGFALGYSSPAIPSLQRAAPP 60

QY 61 APRLDDAAASWFGAVVTLGAAAGVGLGGWLVDRAGKLSLLCSVFFVAGFAVITAAQDV 120
 DB 61 APRLDDAAASWFGAVVTLGAAAGVGLGGWLVDRAGKLSLLCSVFFVAGFAVITAAQDV 120

QY 121 WMLLGRLLTGLACGVASLVAFYISETAYPAVRGLLGCVCVQVMMVVVGLLAYLAGWVLE 180
 DB 121 WMLLGRLLTGLACGVASLVAFYISETAYPAVRGLLGCVCVQVMMVVVGLLAYLAGWVLE 180

QY 181 WRWLAVLGCVPVPSLMLLMCFMPEPRFLLTQHRROEAMAAALRFLWGSEQGWEDDPPIGAE 240
 DB 181 WRWLAVLGCVPVPSLMLLMCFMPEPRFLLTQHRROEAMAAALRFLWGSEQGWEDDPPIGAE 240

QY 241 QSHLALLRQPGIYKPFIIIGVSIIMAFQQLSGVNAVNFYAEITIFEAKFKDSSLASVWVGV 300
 DB 241 QSHLALLRQPGIYKPFIIIGVSIIMAFQQLSGVNAVNFYAEITIFEAKFKDSSLASVWVGV 300

QY 301 IOVLFTVAALIMDRAGRLLLVLSGVVWVFTSFAFGYFKLTQGGPGNSHVAISAPVS 360
 DB 301 IOVLFTVAALIMDRAGRLLLVLSGVVWVFTSFAFGYFKLTQGGPGNSHVAISAPVS 360

QY 361 AQPVDASVGLAWLAVGSMCLFIAGFAVWGPIPWLLMSEIFPLHVKGVAATGICVLTNMLM 420
 DB 361 AQPVDASVGLAWLAVGSMCLFIAGFAVWGPIPWLLMSEIFPLHVKGVAATGICVLTNMLM 420

QY 421 AFLVTKFSSLMVLRPYGAFWLASAFCSFVSLFTLFCVPEYKTKLEQITAHFEGR 477
 DB 421 AFLVTKFSSLMVLRPYGAFWLASAFCSFVSLFTLFCVPEYKTKLEQITAHFEGR 477

RESULT 2

US-09-886-954-1
 Sequence 1, Application US/09886954
 Publication No. US20020038464A1
 GENERAL INFORMATION:
 APPLICANT: Charon, Maureen J.
 APPLICANT: Katz, Ellen B.
 TITLE OF INVENTION: NOVEL GLUCOSE TRANSPORTER/SENSOR PROTEIN AND USES THEREOF
 FILE REFERENCE: 96700/667
 CURRENT APPLICATION NUMBER: US/09/886,954
 CURRENT FILING DATE: 2001-06-21
 NUMBER OF SEQ ID NOS: 1
 SOFTWARE: Patent In version 3.2
 SEQ ID NO 1
 LENGTH: 477
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-886-954-1

Query Match 99.7%; Score 2450; DB 9; Length 477;
 Best Local Similarity 99.8%; Pred. No. 1.2e-202;
 Matches 476; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTPEDPEETQPLGGPPGAPRRVFLAAFAAALGFLSFGFALGYSSPAIPSLQRAAPP 60
 DB 1 MTPEDPEETQPLGGPPGAPRRVFLAAFAAALGFLSFGFALGYSSPAIPSLQRAAPP 60

QY 61 APRLDDAAASWFGAVVTLGAAAGVGLGGWLVDRAGKLSLLCSVFFVAGFAVITAAQDV 120
 DB 61 APRLDDAAASWFGAVVTLGAAAGVGLGGWLVDRAGKLSLLCSVFFVAGFAVITAAQDV 120

QY 121 WMLLGRLLTGLACGVASLVAFYISETAYPAVRGLLGCVCVQVMMVVVGLLAYLAGWVLE 180

Query Match 60.7%; Score 1491; DB 15; Length 326;
 Best Local Similarity 96.3%; Pred. No. 3.8e-120;
 Matches 289; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 164 MVVVGILLAYLAGWLEWRWLAIVGCVPSPSMLLLMCFMPETPRFLLTQHRQEAAMALR 223
 DB 1 MVVVGILLAYLAGWLEWRWLAIVGCVPSPSMLLLMCFMPETPRFLLTQHRQEAAMALR 60

QY 224 FLWGSQGWEDPPIGAEQSFHLALLRPGIYKPIIIGVSLMAFQOLSGVNAVWVFAETIF 283
 DB 61 FLWGSQGWEDPPIGAEQSFHLALLRPGIYKPIIIGVSLMAFQOLSGVNAVWVFAETIF 120

QY 284 BEAKPKDSSLASVVVGVIVLFTVAALIMDRAGRELLLVLSGVVMVFTSFAFGYFKLT 343
 DB 121 BEAKPKDSSLASVVVGVIVLFTVAALIMDRAGRELLLVLSGVVMVFTSFAFGYFKLT 180

QY 344 QGGPGNSSHVASISAPVSAQPVDSVGLAVGSMCLFIAGFAVWGPPIFWLLMSBIFPL 403
 DB 181 QGGPGNSSHVASISAPVSAQPVDSVGLAVGSMCLFIAGFAVWGPPIFWLLMSBIFPL 240

QY 404 HVKGVATGCVLTNMLMAFLVTKFSSLMVLRPYGAFWLASAFCLFSVLTFCVPEPK 463
 DB 241 HVKGVATGCVLTNMLMAFLVTKFSSLMVLRPYGAFWLASAFCLFSVLTFCVPEPK 300

RESULT 4
 US-10-169-395-9
 ; Sequence 9, Application US/10169395
 ; Publication No. US20040034192A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KATO, Seishi
 ; APPLICANT: KIMURA, Tomoko
 ; TITLE OF INVENTION: HUMAN PROTEINS HAVING HYDROPHOBIC DOMAINS AND DNAs ENCODING
 ; TITLE OF INVENTION: THESE PROTEINS
 ; FILE REFERENCE: 01997.015100.US
 ; CURRENT APPLICATION NUMBER: US/10/169,395
 ; PRIOR FILING DATE: 2002-11-29
 ; PRIOR APPLICATION NUMBER: JP 2000-585
 ; PRIOR FILING DATE: 2000-01-06
 ; PRIOR APPLICATION NUMBER: JP 2000-588
 ; PRIOR FILING DATE: 2000-01-06
 ; PRIOR APPLICATION NUMBER: JP 2000-2299
 ; PRIOR FILING DATE: 2000-01-11
 ; PRIOR APPLICATION NUMBER: JP 2000-26862
 ; PRIOR FILING DATE: 2000-02-03
 ; PRIOR APPLICATION NUMBER: JP 2000-58367
 ; PRIOR FILING DATE: 2000-03-03
 ; PRIOR APPLICATION NUMBER: PCT/JP00/09359
 ; PRIOR FILING DATE: 2000-12-28
 ; NUMBER OF SEQ ID NOS: 150
 ; SEQ ID NO 9
 ; LENGTH: 262
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-169-395-9

Query Match 46.2%; Score 1135; DB 15; Length 262;
 Best Local Similarity 91.4%; Pred. No. 1.5e-89;
 Matches 223; Conservative 2; Mismatches 7; Indels 12; Gaps 1;

QY 1 MTPEDPEETQPLIGPPGSAQRGRVFLAAFAAALGPLSFGFALGYSSPAIFSLQRAAPP 60
 DB 1 MTPEDPEETQPLIGPPGSAQRGRVFLAAFAAALGPLSFGFALGYSSPAIFSLQRAAPP 60

QY 61 APRLDDAAASWFGAVVTLGAAAGVVLGGWLVDRAGRKLSELLCSVPFVAGFAVITAAQDV 120
 DB 61 APRLDDAAASWFGAVVTLGAAAGVVLGGWLVDRAGRKLSELLCSVPFVAGFAVITAAQDV 120

QY 121 WWLGGRLTGLACGVASIVAPVYISEIAYPAVRGLLSCVQLMVVVGILLAYLAGWLE 180
 DB 121 WWLGGRLTGLACGVASIVAPVYISEIAYPAVRGLLSCVQLMVVVGILLAYLAGWLE 180

QY 181 WRWLAIVGCVPSPSMLLLMCFMPETPRFLLTQHRQEAAM-----AALRFLMGS 228
 DB 181 WRWLAIVGCVPSPSMLLLMCFMPETPRFLLTQHRQEAAMGLVRCGHGVOHECLRLLQA 240

QY 229 EQGW 232
 DB 241 DPGW 244

RESULT 5
 US-10-157-031-359
 ; Sequence 359, Application US/10157031
 ; Publication No. US20030108890A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baranova, A. V.
 ; APPLICANT: Yankovsky, N. K.
 ; APPLICANT: Kozlov, A. P.
 ; APPLICANT: Lobashev, A. V.
 ; APPLICANT: Krukovskaya, L. L.
 ; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
 ; FILE REFERENCE: 2760-103
 ; CURRENT APPLICATION NUMBER: US/10/157,031
 ; CURRENT FILING DATE: 2002-05-30
 ; NUMBER OF SEQ ID NOS: 415
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 359
 ; LENGTH: 507
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-157-031-359

Query Match 38.6%; Score 948; DB 14; Length 507;
 Best Local Similarity 42.7%; Pred. No. 4.6e-73;
 Matches 216; Conservative 80; Mismatches 166; Indels 44; Gaps 11;

QY 10 QPLIG-----PPGSAQRGRVFLAAFAAALGPLSFGFALGYSSPA 50
 DB 3 EPLLGAEGFDYDFPEKPPSPGDRARVGTLMQKRVFLATFAAVLGNFSGYALVYTSV 62

QY 51 IPSLQRAAPPRLDDAAASWFGAVVTLGAAAGVVLGGWLVDRAGRKLSELLCSVPFVAG 110
 DB 63 IPELERSLDDPHLTKSQASWFGSVTLGAAAGVLSAMLLNDLLKGLSIFMSVPSAAG 122

QY 111 FAVITAAQDVWMLGGRLTGLACGVASIVAPVYISEIAYPAVRGLLSCVQLMVVVGIL 170
 DB 123 YALMAGAGHLMWLLGRTLTGFGGLTAACIPVYVSEIAPPVRCALGATPQLMVAFGSL 182

QY 171 LAYLAGWLEWRWLAIVGCVPSPSMLLLMCFMPETPRFLLTQHRQEAAMALRFLMGSQ 230
 DB 183 SLYALGGLLIPWRWLAIVAGEAPVLIIMLLSFMNSPRFLLSRGRDEALRALAWLRGTDV 242

QY 231 G--WEDPPIG--AEQSFHL--ALLRQGIYKPIIIGVSLMAFQOLSGVNAVWVFAETIF 283
 DB 243 DVHWEFQIQDNVRRQSSRVSAEARAPHVCRPIITVALLMRLQQLTGITPILVYLQSLP 302

QY 284 BEAKF-----KDSLSASVVVGVIVLFTVAALIMDRAGRELLLVLSGVVMVFTSFAFGY 339
 DB 303 DSTAVLLPPKDD---AAIVGAVRLLSVLIAALTMLDLAGRKVLLFVSAAMFAANLTGLY 359

QY 340 FKLTOGGPGNSSHVASISAPV-----AQPVDSVGLAVGSMCLFIAGFAVWGPPI 392
 DB 360 IHF---GPRPLSPNSTAGLESSESGDLAQLAAPAGYLTLPVPLLATMLFMGYAVWGPPI 416

QY 393 PHLLMSEIFPLHKVGVATGCVLTNMLMAFLVTKFSSLMVLRPYGAFWLASAFCLFSV 452
 DB 417 TWLLMSEVLPURARGVAGLCLVLAASLWTAFLVTKSFLPVVSTFGLQVPPFFFAAICLVSL 476

QY 453 LFTLFCVPETKGTLEQITAHEF--GR 477
 DB 477 VFTGCCVETKGRSLEQIESFFRTGR 502

RESULT 6

US-10-755-889-110
 ; Sequence 110, Application US/10755889
 ; Publication No. US20040171823A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bristol-Myers Squibb Company
 ; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
 ; FILE REFERENCE: PATWAY
 ; CURRENT APPLICATION NUMBER: US/10/755,889
 ; PRIOR FILING DATE: 2004-01-13
 ; PRIOR APPLICATION NUMBER: U.S. 60/440,068
 ; PRIOR FILING DATE: 2003-01-14
 ; PRIOR APPLICATION NUMBER: U.S. 60/469,757
 ; PRIOR FILING DATE: 2003-05-12
 ; NUMBER OF SEQ ID NOS: 823
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 110
 ; LENGTH: 507
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-755-889-110

Query Match 38.8%; Score 948; DB 16; Length 507;
 Best Local Similarity 42.7%; Pred. No. 4.6e-73;
 Matches 216; Conservative 80; Mismatches 166; Indels 44; Gaps 11;
 QY 10 QPLG-----PPGSAAPGR-----RVFLAAFAAALGPLSGFALGYSSPA 50
 Db 3 EPLGAGGPDYDFPEKPPSPGDRKRVGTLOKRVFLATFAVLGNFSGYALVYTSV 62
 QY 51 IFSLQRAAPAPRLDDAAASWFGAVVTLGAAAGVGLGWLDVDRAGKLSLLCSCVFPFVAG 110
 Db 63 IPALERSLDPDLHLTKSOASWFSVFTLGAAGGLSAMLNDLLGRKLSIMFSAVPSAAG 122
 QY 111 FAVITAAQDVMWLLGRLTGLACVNASLVAIVYIIEIYPAVRGLLGCSCVQLMVMVVGIL 170
 Db 123 YALMAGAHGLMMLLGRITLTFAGGLTAACIFVYVSEIAPPVGRGALGATPQLMAVFGSL 182
 QY 171 LAYLAGVLEWRLAVLGCVPSPMLLLMCFMPEPRFLLTQHRROEAMAAALRFVWGSQ 230
 Db 183 SLVALGLLPWRWLVAGAPVLIIMLLLSFNPSPRFLLSRGRDEEARLALWLRGIDV 242
 QY 231 G--WEDPPG--AEQFHL--ALLRQPGIYKPFIIIGVSLMAFQQLSGVNAVYFAETIF 283
 Db 243 DVHWEPEIQDNVRRQSSRVSAEAPHCVPITVALIMRLQLQTLGITPILVYLQSI 302
 QY 284 BEAKF---KDSLSASVVGVQLVFTVAALIMDRAGRRLLVLSGVVWVFPSTSAFGAY 339
 Db 303 DSTAVLLPPKDD---AAIVGAVRLLSVLIAALTMDLGRKRVLLFVSAAIMFAANLTLGLY 359
 QY 340 FKLTOGGPNSGHVAISAPVS-----AQPVDASVGLAWLAVGSMCLFIAGFVAGWGP 392
 Db 360 IHP---GPRPLSPNSTATLESWGDLAQPLAAPAGYLTLPVLLATMLFIMGYAVGWGPI 416
 QY 393 PMLLMSFPLHVKGVATGICVLTNWLMAFLVTKFSSLMFVLRPYGAFWLASAPCFISV 452
 Db 417 TWLLMSEVLPRLARGVAGSLCVLWLAFTVLTGKFLPVVSTFGVQVPEFFFAALCLVSL 476
 QY 453 LFTFLCVPTKTKTLEQITAFHE-GR 477
 Db 477 VFTGCCVPTKGRSLEQIESFFRMGR 502

RESULT 7
 US-10-051-909-38
 ; Sequence 38, Application US/10051909
 ; Publication No. US20020199217A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Helentjaris, Tim
 ; APPLICANT: Hertz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott

; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB1163 US CIP
 ; CURRENT APPLICATION NUMBER: US/10/051,909
 ; PRIOR FILING DATE: 2002-01-17
 ; PRIOR APPLICATION NUMBER: 60/083,044
 ; PRIOR FILING DATE: April 24, 1998
 ; NUMBER OF SEQ ID NOS: 38
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 38
 ; LENGTH: 501
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 US-10-051-909-38

Query Match 29.4%; Score 723; DB 13; Length 501;
 Best Local Similarity 35.6%; Pred. No. 1.2e-53;
 Matches 173; Conservative 89; Mismatches 172; Indels 52; Gaps 8;
 QY 17 GGSAPRRRRVFLAFA-----AALGSLSFGEALGYSSPAIPSLORAAPAP 62
 Db 39 GGGGNGSRLGSSAYSRLDSSVAVLCTLIIVALGPIQFGTCGFSPT----QDAIISDL 94
 QY 63 RLDAAASWFGAVVTLGAAAGVGLGWLDVDRAGKLSLLCSCVFPFVAGFVITAAQDVM 122
 Db 95 GLTLESEFLGSLSNVGMVGAIASGQIAEYIGRKSMLIAAIPNIIGLWLAISFAKDSF 154
 QY 123 LLGRLTGLACVNASLVAIVYIIEIYPAVRGLLGCSCVQLMVMVVGILVLAGWLEWR 182
 Db 155 LFMGRLLGEGFVGVISYVYVYVYIABIAIAPQTRGALGSVNLQSVITIGILLVLLGMFV 214
 QY 183 WLAVLGCVPSPMLLLMCFMPEPRFLLTQHRROEAMAAALRFVWGSQGWEDPPIGAB-- 240
 Db 215 ILSVLGILPCSLIPGLFFIPESRFLWAKMGKMGDEFSSLQVLRGF-----TDIAEVN 269
 QY 241 -----QSFHLALRQPGIYKPFIIIGVSLMAFQQLSGVNAVYFAETIFEEAKF 288
 Db 270 EIKRSVQSSRRRTTRFADIKKQVSPVPLMVGIGLLVQLQSGVNLFYAASIFKAAAGL 329
 QY 289 KDSLSASVVGVQLVFTVAALIMDRAGRRLLVLSGVVWVFPSTSAFGAYFKLTQGGPG 348
 Db 330 TNSNLATPGLGVQVAVTGVVTVLTDKAGRRLLIISTGTMTITLVVVSVSF-FVKDNIT 388
 QY 349 NSSHVAISAPVSAPVDASVGLAWLAVGSMCLFIAGFVAGWGPVPLMSEIFPLHVKG 408
 Db 389 NSHL-----YSV-MSMLSLVGLVAVFVIFSLGLGALPWIIMSEILPWNKSL 435
 QY 409 ATGICVLTNWLMAFLVTKFSSLMFVLRPYGAFWLASAPCFISVLFVLPVPTKTKTLE 468
 Db 436 AGSVATLANLWLTAWLITWT-ASLMLSWSNGTFAIYAAVACAGTLVFCVCLWVPTKGR 494
 QY 469 QITAFH 474
 Db 495 EIAFSF 500

RESULT 8
 US-10-424-599-275264
 ; Sequence 275264, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 275264
 ; LENGTH: 484
 ; TYPE: PRT

```

; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_90584C.1.pep
US-10-424-599-275264

```

```

Query Match 29.1%; Score 716; DB 15; Length 484;
Best Local Similarity 35.8%; Pred. No. 4.5e-53;
Matches 164; Conservative 89; Mismatches 174; Indels 34; Gaps 7;

```

```

QY 26 VFLAFAAALGSLGSGFALGYSPPALPSLQRAAPPAPRDLDDAAASWFGAVVTLGAAAGGV 85
Db 45 VLFVCLVALGPIQFGFTGGSPTQDAIIRDL----LTLSEFVSFGLSNVGMVGA 100
QY 86 LGGWLVDRAGRKLSLLCSVPPFVAGFAVITAADVWMLLGGRLTGLACGVASLVAPVYI 145
Db 101 ASQIAEYIGRKGSLMIAAIPNIIGWLAISFAKDSFLYMGRLLEGGFVGIISYVVPVYI 160
QY 146 SEIYPAVRGLLGSVQLMVMVVGILLAYLAGWLEWRWLAIVLGCVPSPMLLMLMCFMPEP 205
Db 161 AEIAPQNLRRGGLSVNQLSVTIGIMLAYLLGLFVNRVLAAILGILPCTVLIPLGFFIPES 220
QY 206 PRFLTQHRROEAMALRFLMGSEOGWEDPPI-----GAEQFHLALLRQPGI 253
Db 221 PRWLAQWMDIPEETSLOVLRGPD--TDISVEHIEKRSVASTGKRAIRFADLKRKY 277
QY 254 YKPFIIIGVSLMAFQQLSGVNAVMEVAETIPEEAKFKDSSLASVYVGVIOVLFATAALIM 313
Db 278 WPLVMVGIILLVLOQLSGINGILFYSTTIFANAGISSSEAAVGLGAVQVIATGISTWLV 337
QY 314 DRAGRLLLVLSGVVWVFTSAPGAYFKLTQGGPNSHVSAIPVSAQPDVDSVGLAW 373
Db 338 DKSGRRLLIISVVNTSLLVSIAYFL-EGVVSDESHL-----PSILGIIVSIVGLVAM 391
QY 374 AVGSMCLFIAGFVAGVGPPIWLLMSEIPLPHVKGVATGICVLTNWLMAFLVTKKESLME 433
Db 392 VI-----GFLSGLGPPIWLLMSEIPLPHVKGVATGICVLTNWLMAFLVTKKESLME 443
QY 434 VLRYPGAFWLASAFICFISVLFVLCVPEKTKTLEQITAHF 474
Db 444 -NSSGGTFIYVVAFTAIAFMWVPEKTKLEIPIQSF 483

```

```

RESULT 9
US-10-437-963-163544
; Sequence 163544, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 163544
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_62530C.1.pep
US-10-437-963-163544

```

```

Query Match 28.6%; Score 702; DB 16; Length 501;
Best Local Similarity 35.9%; Pred. No. 7.5e-52;
Matches 166; Conservative 80; Mismatches 178; Indels 38; Gaps 6;

```

```

QY 27 FLAFAAALGSLGSGFALGYSPPALPSLQRAAPPAPRDLDDAAASWFGAVVTLGAAAGGV 86
Db 85 FLCTLIVALGPIQFGFTGGSPTQDAIIRDL----LTLSEFVSFGLSNVGMVGA 140
QY 87 GGMWLVDRAGRKLSLLCSVPPFVAGFAVITAADVWMLLGGRLTGLACGVASLVAPVYIS 146
Db 141 SQMAEYIGRKGSLMIAAIPNIIGWLAISFAKDSFLYMGRLLEGGFVGIISYVVPVYIA 200
QY 147 EYAPVAVRGLLGSVQLMVMVVGILLAYLAGWLEWRWLAIVLGCVPSPMLLMLMCFMPEP 206
Db 201 EISPNMRGALGSVNQLSVTIGIMLAYLLGLFVNRVLAAILGILPCTVLIPLGFFIPESP 260

```

```

; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_90584C.1.pep
US-10-424-599-275264

```

```

Query Match 29.1%; Score 716; DB 15; Length 484;
Best Local Similarity 35.8%; Pred. No. 4.5e-53;
Matches 164; Conservative 89; Mismatches 174; Indels 34; Gaps 7;

```

```

QY 27 FLAFAAALGSLGSGFALGYSPPALPSLQRAAPPAPRDLDDAAASWFGAVVTLGAAAGGV 86
Db 63 FLCTLIVALGPIQFGFTGGSPTQDAIIRDL----LTLSEFVSFGLSNVGMVGA 118
QY 87 GGMWLVDRAGRKLSLLCSVPPFVAGFAVITAADVWMLLGGRLTGLACGVASLVAPVYIS 146
Db 119 SQMAEYIGRKGSLMIAAIPNIIGWLAISFAKDSFLYMGRLLEGGFVGIISYVVPVYIA 178
QY 147 EYAPVAVRGLLGSVQLMVMVVGILLAYLAGWLEWRWLAIVLGCVPSPMLLMLMCFMPEP 206
Db 179 EISPNMRGALGSVNQLSVTIGIMLAYLLGLFVNRVLAAILGILPCTVLIPLGFFIPESP 238
QY 207 RELLTQHRROEAMALRFLMGSEOGWEDPPIGAE-----OSFHLALLRQPG 252
Db 239 RWLAKWMDIPEETSLOVLRGPD--TDISAEVNDIKRAVASANKRTRIRFOELNOKK 293
QY 253 YKPFIIIGVSLMAFQQLSGVNAVMEVAETIPEEAKFKDSSLASVYVGVIOVLFATAAL 312
Db 294 YRTPILIGLGLVLOQLSGINGILFYSAIFKRAAGLITNSDLATCAIGALQVLAATGVTW 353
QY 313 MDRAQRLLLVLSGVVWVFTSAPGAYFKLTQGGPNSHVSAIPVSAQPDVDSVGLAW 372
Db 354 LDRAGRRILLIISVAGTMSLLAVAVVFL-KDSISQDSHMYT-----LSM 399
QY 373 LAVGSMCLFIAGFVAGVGPPIWLLMSEIPLPHVKGVATGICVLTNWLMAFLVTKKESL 432
Db 400 ISLVALVAFVIAFSGMGAIPIWIMSEIILPVSIKSLAGSFATLANWLTSGITMT-ANLM 458
QY 433 EVLRYPGAFWLASAFICFISVLFVLCVPEKTKTLEQITAHF 474
Db 459 LSWAGGTFVSVVWVSAFTLVFVILWVPEKTKRLEBIQMSF 500

```

```

RESULT 10
US-10-425-114-63431
; Sequence 63431, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 63431
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4371-017-F11_FLI.pep
US-10-425-114-63431

```

```

Query Match 28.6%; Score 702; DB 15; Length 523;
Best Local Similarity 35.9%; Pred. No. 7.9e-52;
Matches 166; Conservative 80; Mismatches 178; Indels 38; Gaps 6;

```

```

QY 27 FLAFAAALGSLGSGFALGYSPPALPSLQRAAPPAPRDLDDAAASWFGAVVTLGAAAGGV 86
Db 85 FLCTLIVALGPIQFGFTGGSPTQDAIIRDL----LTLSEFVSFGLSNVGMVGA 140
QY 87 GGMWLVDRAGRKLSLLCSVPPFVAGFAVITAADVWMLLGGRLTGLACGVASLVAPVYIS 146
Db 141 SQMAEYIGRKGSLMIAAIPNIIGWLAISFAKDSFLYMGRLLEGGFVGIISYVVPVYIA 200
QY 147 EYAPVAVRGLLGSVQLMVMVVGILLAYLAGWLEWRWLAIVLGCVPSPMLLMLMCFMPEP 206
Db 201 EISPNMRGALGSVNQLSVTIGIMLAYLLGLFVNRVLAAILGILPCTVLIPLGFFIPESP 260

```


QY 180 ---EWRWLAVLG--CVPPSLMILLMCMFMPETPRFLLTQHRROBAMAALRFLMGSEQ-GWE 233
 Db 179 ALNSGWRIPGLQLVLPALLLIGLLFLPESPRWLVEKLEAREVLAKLGRVEDVDQE 238
 QY 234 DPPIGAEOQSFHLA-----LLR---QPGIYKPFIIIGYSLMAFOQLSGVNAVMYFA 279
 Db 239 IQEIKAELEATVSEKAKGASWEGELFRGTRPAVRQRLMLGVMWLOAFOQLTGAINAIFYYS 298
 QY 280 ETIFEEAKEFKD---SSLASVVVGVIOVLEFTAVAL-IMDRAGRRLLLVLSGVVWFSTSA 335
 Db 299 PFIKSVGVSDSVASLLVTIIVGVVNFVFTVALIFLVDVFRGRRPRLLLGAAAGMAICFLI 358
 QY 336 FGAYFKLTQGGPGNSHVAISAPVSAQPDVDSVGLAWL-AVGSMLCFIAGFVAGWGPWPW 394
 Db 359 LGA-----SIGVALLLNKPKDPSKKAAGIVAIIVFILLFIAFFALGWPWPW 405
 QY 395 LLMSEIFPLHKVGVATGICVLTNWLMAFLVTKFSSLMELVRPY--GAFWLA---SAFCI 449
 Db 406 VILSELFPKVRKSKALALATAANWLANFII-----GFLFPYITGAIGLALGGVVFIV 457
 QY 450 FS-----VLFILFCVPEKTKLEQITAHF 474
 Db 458 FAGLLVLFILVFFVFPVETKGRTEEBEELF 488

RESULT 15
 US-10-094-059-4 ;
 ; Sequence 4, Application US/10094059
 ; Publication No. US20020127650A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Millennium Pharmaceuticals Inc.
 ; TITLE OF INVENTION: 32468, A Human Sugar Transporter Family Member and
 ; TITLE OF INVENTION: Uses Therefor
 ; FILE REFERENCE: MPIO1-040PIRM
 ; CURRENT APPLICATION NUMBER: US/10/094,059
 ; PRIOR FILING DATE: 2002-03-08
 ; PRIOR APPLICATION NUMBER: 60/275,053
 ; PRIOR FILING DATE: 2001-03-12
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 488
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: consensus
 US-10-094-059-4

Query Match 25.0%; Score 615; DB 13; Length 488;
 Best Local Similarity 35.0%; Pred No. 2.3e-44;
 Matches 179; Conservative 74; Mismatches 168; Indels 90; Gaps 18;
 QY 30 AFAAALGPLSFGFALGYSSPAIP-----SLQRAAPPAPRLDDAASWFGAVVT 77
 Db 2 ALVAALGG--GFLRGYDGTGVIIGGLFALIDFLFRGLLTSGLAELVGYSTVLTGLVVS 58
 QY 78 ---LGAAGVGLGWLVDVDRGKLSLLCSVPFVAGFAVITAAQ-----DWW---MLGG 126
 Db 59 IFFLGLIGSLFAGLGRFRGRKSLIALLVLFVIGALLSAAAGPYTTIGLMAFYLLIVG 118
 QY 127 RLTLGLACVSLVAVPVYISEIYAVRGLGSCVQLMVMVVGILLAYLAGWVL-----179
 Db 119 RVLVGLVGGASVIVPMYIASEIAPKALRGALGSLYQLAITIGILVAAIIGLGNKTNDS 178
 QY 180 ---EWRWLAVLG--CVPPSLMILLMCMFMPETPRFLLTQHRROBAMAALRFLMGSEQ-GWE 233
 Db 179 ALNSGWRIPGLQLVLPALLLIGLLFLPESPRWLVEKLEAREVLAKLGRVEDVDQE 238
 QY 234 DPPIGAEOQSFHLA-----LLR---QPGIYKPFIIIGYSLMAFOQLSGVNAVMYFA 279
 Db 239 IQEIKAELEATVSEKAKGASWEGELFRGTRPAVRQRLMLGVMWLOAFOQLTGAINAIFYYS 298

QY 280 ETIFEEAKEFKD---SSLASVVVGVIOVLEFTAVAL-IMDRAGRRLLLVLSGVVWFSTSA 335
 Db 299 PFIKSVGVSDSVASLLVTIIVGVVNFVFTVALIFLVDVFRGRRPRLLLGAAAGMAICFLI 358
 QY 336 FGAYFKLTQGGPGNSHVAISAPVSAQPDVDSVGLAWL-AVGSMLCFIAGFVAGWGPWPW 394
 Db 359 LGA-----SIGVALLLNKPKDPSKKAAGIVAIIVFILLFIAFFALGWPWPW 405
 QY 395 LLMSEIFPLHKVGVATGICVLTNWLMAFLVTKFSSLMELVRPY--GAFWLA---SAFCI 449
 Db 406 VILSELFPKVRKSKALALATAANWLANFII-----GFLFPYITGAIGLALGGVVFIV 457
 QY 450 FS-----VLFILFCVPEKTKLEQITAHF 474
 Db 458 FAGLLVLFILVFFVFPVETKGRTEEBEELF 488

Search completed: February 24, 2005, 06:04:07
 Job time : 134 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 05:44:20 ; Search time 44 Seconds
(without alignments)
1043.077 Million cell updates/sec

Title: US-09-886-954A-1

Perfect score: 2457

Sequence: 1 MTPPEPEPQLPPGGSA.....CVPETKGTLEQITAHPEGR 477

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 483416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Table with columns: Result No., Score, Query Match, Length, DB ID, Description. Lists various protein entries like 'hypothetical prote', 'probable sugar tra', 'metabolite transpo', etc.

Table with columns: 30, 542, 22.1, 534, 2, S38435, hexose transport p, etc. Lists various protein entries and their scores.

ALIGNMENTS

RESULT 1

E96782
hypothetical protein F22H5.6 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C:Accession: E96782

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.F.; Hughes, B.; Huijar, L.
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: E96782

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-487 <STO>

A:Cross-references: UNIPROT:Q9FRL3; GB:AE005173; NID:g10092276; PIDN:AAG12689.1; GSPDB:

C:Genetics:

A:Gene: F22H5.6

A:Map position: 1

C:Superfamily: glucose transport protein

Query Match 29.9%; Score 734; DB 2; Length 487;
Best Local Similarity 37.5%; Pred. No. 1.1e-45;
Matches 173; Conservative 77; Mismatches 177; Indels 34; Gaps 6;

OY 26 VFLAFAAALGFLSFGALGYSIPAIPSLQRAAPAPRLDDAAASHWFGAVVILGAAAGGV 85

Db 48 VLACVLIIVALGPIQFGFGYSSPT---QAATKDLGLTVEYSYVFGSLNVMGAVI 103

OY 86 LGGWLVDRAGKRLSLLGCSVRFVAGFAVITAAQDVVWMLGGRLTLTLAGVSLVAPVVI 145

Db 104 ASQQAETIYGRKSLMIRAIPIIIGWLCISFAKDFLYMGRLLGEGFGVGIISYTPVVI 163

OY 146 SEIAYPAVRGLGSCVQLMVMVVGILLAYLAGVLEWRWLVAVLGVPPSLMLLMCFMPE 205

Db 164 AEIAQNMREGGLSVNQLSVTIGIMLAYLGLFVPRWILAVLGLPCTLLIPLGLPIPE 223

OY 206 PRLFTQHRQRQMAALRFLMGSEQGWEDPPIGAEQ-----SFHLALLRQPGI 253

Db 224 PRLARMGMTDFEFTSLQVLRGFE---TDITVEVNEIKRSVASSTKRNTRVFRVFLKRRY 280

OY 254 YKPFITIGVSLMAFQQLSGVNAVYAEITFEBAKFKDSSLASVWGVIVLFTVAALJM 313

Db 281 YPFLMVGVLIVLQJGGINGVLFVSSITFESAGVTSNATFGVAGIQQVATAITWLV 340

OY 314 DRAGRRLLVLSGVVMVFSTSAFAGYFKLTGGPGNSHVAISAPVSAQFVDASVGLAWL 373

Db 341 DRAGRRLLTISVSGWMTSLVVAANFYLKE-----FVSDPS-----DMYSLWLSIL 386
 QY 374 AVGSMCLFIAGFAVGGPIPWLLMSBIEFPLHVKGVATGICVLTNMLMAFLVTKFSSLSME 433
 Db 387 SVVGVVAVVVFSLGMPPIWLLMSBIEFPLHVKGVATGICVLTNMLMAFLVTKFSSLSME 445
 QY 434 VLRPYGAFWASAFICFISVLFVLCVPEYKTKLQITAHF 474
 Db 446 AWSGSGTFLYGLVCAFTVVFVTLWVWPETKGRKLEELQSLF 486

RESULT 2
 T14545
 C:Species: probable sugar transporter protein - beet
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
 C:Accession: T14545
 R;Chou, T.J.; Bush, D.R.
 Plant Physiol. 110, 511-520, 1996
 A;Title: Molecular cloning, immunohemical localization to the vacuole, and expression
 A;Reference number: Z18131; MUID:963511183; PMID:8742332
 A;Accession: T14545
 C:Superfamily: glucose transport protein
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-490 <CHI>
 A;Cross-references: UNIPROT:Q39416; EMBL:U43629; NID:gl209755; PIDN:AAB53155.1; PID:gl20
 A;Experimental source: tonoplast
 C;Superfamily: glucose transport protein
 C;Keywords: transmembrane protein

Query Match 29.4%; Score 722; DB 2; Length 490;
 Best Local Similarity 36.1%; Pred. No. 8.1e-45;
 Matches 167; Conservative 81; Mismatches 177; Indels 38; Gaps 5;
 QY 26 VFLAFAAALGPLSGFALGYSPPALPSLQRAAPPAPRLDDAAASWFGAVVTLGAAAGV 85
 Db 50 VLACVLVALGFIQFOFTAGYSPT---QSALTNELGSLVAEYENFSGLSNVGAVGAI 105
 QY 86 LGGLVDRAGRLKSLLLCSVFPVAGFAVITAAQDVMWLLGRLLLGLACVNASLVAPVY 145
 Db 106 ASGOISEYIGRGLMIAAIPNIGLWLAISFAKDSFLYMRGLBGFVGIISYTPVYI 165
 QY 146 SETAYPAVRLGSLCVLMVWVIGILLAYLAGVLEWRWLVAVLGCYPPSPMLLLMCFMPE 205
 Db 166 SETAPQNLKRGALGVNQLSVTIGIMLSYMLGLFVPRWRLAVLIGLIPCTLILPGLFFPES 225
 QY 206 PRELLTQHRROEAMALRFLWGSEOGWEDPPICAE-----DTPDLSLVNEIKRVSASSKRTIRFAELRQ 251
 Db 226 PRLAKWGMEEPEVSLQVLRGF----DTDLSLVNEIKRVSASSKRTIRFAELRQ 280
 QY 252 GIYKPIIIGVSLMAFQOLSGVNAVMPYAEITIEEAKFKDSSLASVVVGVIVQLFTVAAL 311
 Db 281 RYWLPLMIGNLILQOLSGINGLVFYSSTIPEKAGVTSNAATFGLGAVQVIAVVTWTT 340
 QY 312 IMDRAGRLLLVLSGVVWVFSFAFGAYFKLTQGGPNSHVAISAPVSAQPVDAVGLA 371
 Db 341 LVDKSGRLLLVSSSGMTLSLIVVAMSFFLKMVSDSTWYSV-----FS 386
 QY 372 WLAGVSMCLFIAGFAVGGPIPWLLMSBIEFPLHVKGVATGICVLTNMLMAFLVTKFSSLS 431
 Db 387 ILSVGVVAVVVFSLGMPPIWLLMSBIEFPLHVKGVATGICVLTNMLMAFLVTKFSSLS 445
 QY 432 MEVLRPYGAFWASAFICFISVLFVLCVPEYKTKLQITAHF 474
 Db 446 MLSWNSGGTFSIYMWVCAFTVAVFVWVWPETKGRKLEELQSWF 488

RESULT 3
 D70073
 C:Species: metabolite transport protein homolog yxkC - Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C:Accession: D70073
 R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
 A.; Bron, S.; Brouillette, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
 C.; Ehrlich, S.D.; Emmerman, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fuma, S.; Galizzi, A.; Galler
 Iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
 A.; Ogasawara, N.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Y. M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
 A.; Authors: Schleicher, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, K.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yamamoto, K.; Yata, K.; Yoshida, K.
 A;Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AB0868
 R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 A;Experimental source: strain 168
 A;Accession: T14545
 C;Genetics: yxkC
 C;Superfamily: glucose transport protein

Query Match 24.7%; Score 607; DB 2; Length 461;
 Best Local Similarity 31.1%; Pred. No. 1.6e-36;
 Matches 146; Conservative 85; Mismatches 192; Indels 46; Gaps 7;
 QY 25 RVFLAFAAALGPLSGFALGYSPPALPSLQRAAPPAPRLDDAAASWFGAVVTLGAAAGG 84
 Db 6 RXYMIFYFAGLGGLYGDTGIVISGALLFNNDIPLTLTLEGLVVS---MLLGLALFGS 61
 QY 85 VLGLVDRAGRLKSLLLCSVFPVAGFAVITAAQDVMWLLGRLLLGLACVNASLVAPVY 144
 Db 62 ALSGTCSDRWGRKRVVFLSIFIIIGALACAFSQTIGMLIASRVILGLAVGGSTALVPVY 121
 QY 145 ISEIYPAVRLGSLCVLMVWVIGILLAYLAGVLEWRWLVAVLGCYPPSPMLLLMCLM 200
 Db 122 LSEMAPTKIRGLTGMNMLVITGILLAYVIVLFFPEAWRMVWVGLAAVPAVLLIGIA 181
 QY 201 FMPETPRLTQHRROEAMALRFLWGSEOGWEDPPICAEQSFHALLRQ 250
 Db 182 FMPESPRLVWKRGEERARRINWTHDPKDIEMELAEKQGEAE----KKEITLGLVJKA 236
 QY 251 GIYKPIIIGVSLMAFQOLSGVNAVMPYAEITIEEAKFKDSSLASVVVGVIVQLFTAV 308
 Db 237 KWIRPMLLIGVGLAIFQQAVGINTVIYVYPTFTKAGLGTSSASALGTMGILNVMCIT 296
 QY 309 AALIMDRAGRLLLVLSGVVWVFSFAFGAYFKLTQGGPNSHVAISAPVSAQPVDAV 368
 Db 297 AMILIDRVRKRLIWSVGVITLSAALSGLV-LTLG-----LSA 335
 QY 369 GLAVLAVGSMCLFIAGFAVGGPIPWLLMSBIEFPLHVKGVATGICVLTNMLMAFLVTKF 428
 Db 336 STAMVTVVFLGVIVVQAVTQVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWV 395
 QY 429 SSLMEVLRPYGAFWASAFICFISVLFVLCVPEYKTKLQITAHF 477
 Db 396 PLMLSAMGIAWVWVFSVCLLSFFAFYFVWVWVWVWVWVWVWVWVWVWVWVWVWVWV 444

RESULT 4
 AB0868
 L-arabinose isomerase [imported] - Salmonella enterica subsp. enterica serovar Typhi (st
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A;Note: this species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AB0868
 R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

Db 121 VLGIAGVIASTYAPLYLSEMASENVGRKMSYQMLVWTLGIVLAFSLDTSFYSYGNWRAM 180
 Qy 187 LG--CVPPSLMLLMLCFMPEPRFLLTQHRROEAANAARFLWNGSEQWEDPPIGAEQSFH 244
 Db 181 LGLVALPAVLLIILVVFLNSPRWLAEKGRHTEAEVLRMLRDTSEKAREELNEIRESLK 240
 Qy 245 L-----ALLR-QPGIVKPIIGVSLMAFOQLSGVNAVVFYABTIFEEAKF---KDSLSLAS 295
 Db 241 LKQGGWALFKINRVRRAVFLGMLLQAMQOFTGMNLIIMYAPRIFKMGAFITTEOQMIAT 300
 Qy 296 VVVGVIOVLTVAVALIMDRAGRRLLVLSGVVVFSTSAFGAYFKLTQGGPGNSSHVAI 355
 Db 301 LVVGLTFMFATFIAVFTVDKAGKPKALKIGFVSMALGLTLVLG--YCLMDFDNGT----- 352
 Qy 356 SAPVSAQVDASVGLAWLAVGSMCLFIAGFAVWGQPIFWLLMSEIFPLHVKGVATGICVL 415
 Db 353 -----ASSGLSWLSVGMWMTMCIAAGTFLWYALNIAFVGIITFWLLIPETKKNVLEHI 457
 Qy 416 TNWLMAFVTKFSSLSMVEVLRPVGAFWLASAFICFISVLFVFCVPETKGTLEQI 470
 Db 403 TNWVSNMIIIGATFLTLSDSIAAGTFLWYALNIAFVGIITFWLLIPETKKNVLEHI 457

RESULT 7
 E85936
 L-arabinose isomerase (EC 5.3.1.4) - Escherichia coli (strain O157:H7, substrain EDL933)
 C;Species: Escherichia coli
 C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C;Accession: E85936
 R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A;Reference number: A85480; MUID:21074935; PMID:11206551
 A;Accession: E85936
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-472 <STO>
 A;Cross-references: UNIPROT:P09830; GB:AE005174; NID:G12517333; PIDN:AGS7953.1; GSPDB:C
 A;Experimental source: strain O157:H7, substrain EDL933
 C;Genetics:
 C;Superfamily: glucose transport protein
 C;Keywords: intramolecular oxidoreductase; isomerase

Query Match 24.5%; Score 603; DB 2; Length 472;
 Best Local Similarity 33.3%; Pred. No. 3.1e-36;
 Matches 158; Conservative 84; Mismatches 197; Indels 36; Gaps 9;
 Qy 9 TOPLLGPPGSAAPRRVFLAAFAAALGFLSFGALGYSPPAIPSLQRAAPPAPRLDDAA 68
 Db 6 TESALTFRSLRDRTRMMFVSV-AAVAGLLFGLDIDGVIAGALPFTIDHFLVTSRLOE-- 62
 Qy 69 ASWFGAVVTLGAAGVGLGWVDRAGKLSLLLCVFPVAVGAVITAAQDVMWLLGGRL 128
 Db 63 --VWVSSMMLGAAIGALFNGLSFRGKYSLMAGAILFVLSIGSAFATVSEMLIAARV 120
 Qy 129 LTGLACGVASLAVPVISETIAYPAVRELLGSCVOLMVVVGILLIAYLAGVLEW--RWLAV 186
 Db 121 VLGIAGVIASTYAPLYLSEMASENVGRKMSYQMLVWTLGIVLAFSLDTSFYSYGNWRAM 180
 Qy 187 LG--CVPPSLMLLMLCFMPEPRFLLTQHRROEAANAARFLWNGSEQWEDPPIGAEQSFH 244
 Db 181 LGLVALPAVLLIILVVFLNSPRWLAEKGRHTEAEVLRMLRDTSEKAREELNEIRESLK 240
 Qy 245 L-----ALLR-QPGIVKPIIGVSLMAFOQLSGVNAVVFYABTIFEEAKF---KDSLSLAS 295
 Db 241 LKQGGWALFKINRVRRAVFLGMLLQAMQOFTGMNLIIMYAPRIFKMGAFITTEOQMIAT 300
 Qy 296 VVVGVIOVLTVAVALIMDRAGRRLLVLSGVVVFSTSAFGAYFKLTQGGPGNSSHVAI 355
 Db 301 LVVGLTFMFATFIAVFTVDKAGKPKALKIGFVSMALGLTLVLG--YCLMDFDNGT----- 352

Qy 356 SAPVSAQVDASVGLAWLAVGSMCLFIAGFAVWGQPIFWLLMSEIFPLHVKGVATGICVL 415
 Db 353 -----ASSGLSWLSVGMWMTMCIAAGTFLWYALNIAFVGIITFWLLIPETKKNVLEHI 457
 Qy 416 TNWLMAFVTKFSSLSMVEVLRPVGAFWLASAFICFISVLFVFCVPETKGTLEQI 470
 Db 403 TNWVSNMIIIGATFLTLSDSIAAGTFLWYALNIAFVGIITFWLLIPETKKNVLEHI 457

RESULT 8
 B26430
 L-arabinose isomerase (EC 5.3.1.4) - Escherichia coli (strain K-12)
 C;Species: Escherichia coli
 C;Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 09-Jul-2004
 C;Accession: B26430; I40996; I40996; B65067
 R;Maiden, M.C.J.; Davis, E.O.; Baldwin, S.A.; Moore, D.C.M.; Henderson, P.J.F.
 Nature 325, 641-643, 1987
 A;Title: Mammalian and bacterial sugar transport proteins are homologous.
 A;Reference number: A93389; MUID:87115869; PMID:3543693
 A;Accession: B26430
 A;Status: nucleic acid sequence not shown; not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 1-472 <MAI>
 A;Cross-references: UNIPROT:P09830
 R;Maiden, M.C.J.; Jones-Mortimer, M.C.; Henderson, P.J.F.
 J. Biol. Chem. 263, 8003-8010, 1988
 A;Title: The cloning, DNA sequence, and overexpression of the gene araE coding for arabi
 A;Reference number: A28075; MUID:98228015; PMID:2836407
 A;Accession: A28075
 A;Molecule type: DNA
 A;Residues: 1-472 <MA2>
 A;Cross-references: GB:J03732; NID:G145320; PIDN:AAA23469.1; PID:G145321
 R;Stoner, C.; Schleif, R.
 J. Mol. Biol. 171, 369-381, 1983
 A;Title: The araE low affinity L-arabinose transport promoter. Cloning, sequence, trans
 A;Reference number: I40996; MUID:84114868; PMID:6319708
 A;Accession: I40996
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-25, 'Y', 349 'R', <RES>
 A;Cross-references: EMBL:X00272; NID:G40940; PIDN:CAA25075.1; PID:G40941
 R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 .A.; Rose, D.J.; Mau, B.; Shaoh, Y.
 Science 277, 1453-1462, 1997
 A;Title: The complete genome sequence of Escherichia coli K-12.
 A;Reference number: A64720; MUID:97426617; PMID:9278503
 A;Accession: B65067
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-472 <BLAT>
 A;Cross-references: GB:AR000368; GB:U00096; NID:G2367165; PIDN:AACT5880.1; PID:G1789207;
 A;Experimental source: strain K-12, substrain MGL1655
 C;Genetics:
 A;Gene: araE
 A;Map position: 61 min
 C;Superfamily: glucose transport protein
 C;Keywords: arabinose transport; intramolecular oxidoreductase; isomerase; membrane prot

Query Match 24.5%; Score 603; DB 2; Length 472;
 Best Local Similarity 33.3%; Pred. No. 3.1e-36;
 Matches 158; Conservative 84; Mismatches 197; Indels 36; Gaps 9;
 Qy 9 TOPLLGPPGSAAPRRVFLAAFAAALGFLSFGALGYSPPAIPSLQRAAPPAPRLDDAA 68
 Db 6 TESALTFRSLRDRTRMMFVSV-AAVAGLLFGLDIDGVIAGALPFTIDHFLVTSRLOE-- 62
 Qy 69 ASWFGAVVTLGAAGVGLGWVDRAGKLSLLLCVFPVAVGAVITAAQDVMWLLGGRL 128
 Db 63 --VWVSSMMLGAAIGALFNGLSFRGKYSLMAGAILFVLSIGSAFATVSEMLIAARV 120
 Qy 129 LTGLACGVASLAVPVISETIAYPAVRELLGSCVOLMVVVGILLIAYLAGVLEW--RWLAV 186
 Db 121 VLGIAGVIASTYAPLYLSEMASENVGRKMSYQMLVWTLGIVLAFSLDTSFYSYGNWRAM 180

187 LG--CVPPSLMMLLFCMPETPRFLLTQHRROEAMAAALRFLWMSQEQWEDDPPIGAPQSPH 244
 181 LGVLLPAVLLIIVLVPNSPRLAEKGRHIEAEVRLRMLRDTSEKAREELNEIRESLK 240
 245 L-----ALLR-QPGIYKPIIIGVSLMAFOOLSGVNAVVFYAEITFEAKP---KSSLAS 295
 241 LKGGGKALPKINRNRVAVFLGMLLQAMQOFTGMNIIIMYAPRIFKMWAGFTTTEQQMIAT 300
 296 VVGVQVQLVFAAALIMDRAGRRLLLVLSGVVVFSTSAFGAYFKLTOGGFGNSSHVAI 355
 301 LVVGLTFMFAATFAVFDVDAKGRKPAKIGFVMAIGTLVIG--YCLMFDNGT----- 352
 356 SAPVSAQPDVADSYGLAWLAVGSMCLFIAGFVAGWGPWPMLMSEIPLPHVKGVATGICVL 415
 353 -----ASSGLSWSLWMTMCMCIAGYAMSAAPVWMLCSEIQPLKCRDFGICSTT 402
 416 TNLWMAFLVTKFESSLMEVLRPYGAFWLASAFICFVSLFTLFCVPTKGTLEQITAH 470
 403 TNWVSNMIIGATEPLTLLDSIGRAGTFWLYTALNIAFVGTIFWMLIPETKNTLLEHI 457

RESULT 9
 P65079
 galactose-proton symport (galactose transporter) - Escherichia coli (strain K-12)
 C:Species: Escherichia coli
 C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
 C:Accession: F65079
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: F65079
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-464 <BLAT>
 A:Cross-references: UNIPROT:P37021; GB:AE000377; GB:U00096; NID:g2367178; PIDN:AACT5980.
 A:Experimental source: strain K-12, substrain MG1655
 C:Genetics:
 C:Superfamily: glucose transport protein

Query Match 24.1%; Score 591.5; DB 2; Length 464;
 Best Local Similarity 32.4%; Pred. No. 2.1e-35;
 Matches 155; Conservative 83; Mismatches 175; Indels 65; Gaps 11;
 27 FLAFAAALGPLSFGFALGYSIPAIFSLQRAAPPAPRLDDAAASWFGAVVTIGAAAGVGL 86
 16 FFVCFLAALAGLLFGLDIGVIAAGALPPIADEF---QITSHQEQWVSSMMFGAAVAVG 71
 87 GGVLDVDRAGKLSLLCSVFPVAGFAVITAAQDVMMLGGRLLTGLACGVASIVAPVYS 146
 72 SGWLSFKLGRKSLMIGAILFVAGSLFSAAPNVEVLLSRVLLGLAVGASVYTABLYLS 131
 147 ETAYPAVAVRGLGSCVQLVWVVGILLAYLA---GWVLEWRWLAVLGCVPPSLMMLLFCM 202
 132 ETAPEKIRGSMISMYQLMITIGLGYLSDTAFSYGAWRMLGVIIIPAILLLIGVFFL 191
 203 PETPPELLTQHRROEAMAAAL-----RFLWMSQEQWEDDPPIGAPQSPH 244
 192 PDSRPFPAKRRFVDAERVLLRLRDTSAEKRELDIRESLQVKQSGW----- 239
 245 LALLRQPGIYKPII-IGVSLMAFOOLSGVNAVVFYAEITFEAKFKDSS---LASVVVGV 300
 240 -ALFKENSNFRRAVFLGVLQVQOFTGMNIIIMYAPKIFELAGYNTTTEQQMGTIVGL 298
 301 IQVLFATAVALIMDRAGRRLLLVLSGVVVFSTSAFGAYFKLTOGGFGNSHVAISAPVS 360
 299 TNVLAFFIAGLVDRWRKPTLTLGFLVM-----AAG-----MGVLTGMHMHGIHSP-S 346
 361 AQPVDASVGLAWLAVGSMCLFIAGFVAGWGPWPMLMSEIPLPHVKGVATGICVLTNWL 420
 377 ETAPEKIRGSMISMYQLMITIGLGYLSDTAFSYGAWRMLGVIIIPAILLLIGVFFL 191
 203 PETPPELLTQHRROEAMAAAL-----RFLWMSQEQWEDDPPIGAPQSPH 244
 192 PDSRPFPAKRRFVDAERVLLRLRDTSAEKRELDIRESLQVKQSGW----- 239
 245 LALLRQPGIYKPII-IGVSLMAFOOLSGVNAVVFYAEITFEAKFKDSS---LASVVVGV 300
 240 -ALFKENSNFRRAVFLGVLQVQOFTGMNIIIMYAPKIFELAGYNTTTEQQMGTIVGL 298
 301 IQVLFATAVALIMDRAGRRLLLVLSGVVVFSTSAFGAYFKLTOGGFGNSHVAISAPVS 360
 299 TNVLAFFIAGLVDRWRKPTLTLGFLVM-----AAG-----MGVLTGMHMHGIHSP-S 346
 361 AQPVDASVGLAWLAVGSMCLFIAGFVAGWGPWPMLMSEIPLPHVKGVATGICVLTNWL 420

347 AQ-----YFAIAMLMLFVGFAMSGAPLIWLCSEIQPLKGRDFGICSTATNWA 397
 421 AFLVTKFESSLMEVLRPYGAFWLASAFICFVSLFTLFCVPTKGTLEQITAH-FEGR 477
 398 NMIVGATFTLMLTNGNANTFWYAAALNVLFIILLVLPETKHSLEHIERLMLMGR 455

RESULT 10
 C91106
 galactose-proton symport of transport system Ecs3819 [imported] - Escherichia coli (str
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
 C:Accession: C91106
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Iehii, K.; Yokoyama, K.; Han, C.G.
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: C91106
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-464 <HAY>
 A:Cross-references: UNIPROT:Q8XCW7; GB:BA000007; PIDN:BA837242.1; PID:g13363291; GSPDB:
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 C:Superfamily: glucose transport protein

Query Match 23.9%; Score 587.5; DB 2; Length 464;
 Best Local Similarity 32.2%; Pred. No. 4.1e-35;
 Matches 154; Conservative 83; Mismatches 176; Indels 65; Gaps 11;
 27 FLAFAAALGPLSFGFALGYSIPAIFSLQRAAPPAPRLDDAAASWFGAVVTIGAAAGVGL 86
 16 FFVCFLAALAGLLFGLDIGVIAAGALPPIADEF---QITSHQEQWVSSMMFGAAVAVG 71
 87 GGVLDVDRAGKLSLLCSVFPVAGFAVITAAQDVMMLGGRLLTGLACGVASIVAPVYS 146
 72 SGWLSFKLGRKSLMIGAILFVAGSLFSAAPNVEVLLSRVLLGLAVGASVYTABLYLS 131
 147 ETAYPAVAVRGLGSCVQLVWVVGILLAYLA---GWVLEWRWLAVLGCVPPSLMMLLFCM 202
 132 ETAPEKIRGSMISMYQLMITIGLGYLSDTAFSYGAWRMLGVIIIPAILLLIGVFFL 191
 203 PETPPELLTQHRROEAMAAAL-----RFLWMSQEQWEDDPPIGAPQSPH 244
 192 PDSRPFPAKRRFVDAERVLLRLRDTSAEKRELDIRESLQVKQSGW----- 239
 245 LALLRQPGIYKPII-IGVSLMAFOOLSGVNAVVFYAEITFEAKFKDSS---LASVVVGV 300
 240 -ALFKENSNFRRAVFLGVLQVQOFTGMNIIIMYAPKIFELAGYNTTTEQQMGTIVGL 298
 301 IQVLFATAVALIMDRAGRRLLLVLSGVVVFSTSAFGAYFKLTOGGFGNSHVAISAPVS 360
 299 TNVLAFFIAGLVDRWRKPTLTLGFLVM-----AAG-----MGVLTGMHMHGIHSP-S 346
 361 AQPVDASVGLAWLAVGSMCLFIAGFVAGWGPWPMLMSEIPLPHVKGVATGICVLTNWL 420
 377 ETAPEKIRGSMISMYQLMITIGLGYLSDTAFSYGAWRMLGVIIIPAILLLIGVFFL 191
 203 PETPPELLTQHRROEAMAAAL-----RFLWMSQEQWEDDPPIGAPQSPH 244
 192 PDSRPFPAKRRFVDAERVLLRLRDTSAEKRELDIRESLQVKQSGW----- 239
 245 LALLRQPGIYKPII-IGVSLMAFOOLSGVNAVVFYAEITFEAKFKDSS---LASVVVGV 300
 240 -ALFKENSNFRRAVFLGVLQVQOFTGMNIIIMYAPKIFELAGYNTTTEQQMGTIVGL 298
 301 IQVLFATAVALIMDRAGRRLLLVLSGVVVFSTSAFGAYFKLTOGGFGNSHVAISAPVS 360
 299 TNVLAFFIAGLVDRWRKPTLTLGFLVM-----AAG-----MGVLTGMHMHGIHSP-S 346

RESULT 11
 F85951
 galactose-proton symport of transport system [imported] - Escherichia coli (strain O157
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C:Accession: F85951
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhe
 iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca

Db 434 NWISNLIVAQTELLTIAAAGTGMTELLLAGIAVLAVIFVIVFVPEQGLTFSEV 487

Search completed: February 24, 2005, 05:53:46
Job time : 46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 05:43:22 ; Search time 180 Seconds
(without alignments)
1357.010 Million cell updates/sec

Title: US-09-886-954A-1
Perfect score: 2457
Sequence: 1 MTPPEPTEPQLLPPGGSA.....CVPETKTKLEITAHFEGR 477

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Table with columns: Result No., Score, Query Match Length, DB ID, Description. Contains 31 rows of search results.

Table with columns: ID, Q8WUZ9, PRELIMINARY, PRT, 477 AA. Lists various protein entries and their identifiers.

ALIGNMENTS

Table with columns: RESULT 1, O8WUZ9, PRELIMINARY, PRT, 477 AA. Shows sequence alignments for various proteins.

Binds cytochalasin B in a glucose-inhibitable manner. Seems to be a dual-specific sugar transporter as it is inhibitable by fructose (By similarity).

-1- SUBCELLULAR LOCATION: Integral membrane protein. Principally intracellular. May move between intracellular vesicles and the plasma membrane. The dileucine internalization motif is critical for intracellular sequestration (By similarity).

-1- TISSUE SPECIFICITY: Highly expressed in testis, but not in testicular carcinoma. Lower amounts present in most other tissues.

-1- INDUCTION: In testis, down-regulated by estrogen.

-1- SIMILARITY: Belongs to the sugar transporter family. Glucose transporter subfamily.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch)

CC	EMBL; Y17801; CAB89809.1; ..	DR	EMBL; Y17801; CAB89809.1; ..	1	25	Cytoplasmic (Potential).
CC	EMBL; AJ245937; CAB75702.1; ..	DR	EMBL; AJ245937; CAB75702.1; ..	1	46	1 (Potential).
CC	Gene; HGNC:13812; SLC2A8.	DR	Gene; HGNC:13812; SLC2A8.	26	70	Extracellular (Potential).
CC	H-InvDB; HIK0008396; ..	DR	H-InvDB; HIK0008396; ..	47	91	2 (Potential).
CC	MIM; 605245; ..	DR	MIM; 605245; ..	71	96	Cytoplasmic (Potential).
CC	GO; GO:0005887; C:integral to plasma membrane; TAS.	DR	GO; GO:0005887; C:integral to plasma membrane; TAS.	92	96	3 (Potential).
CC	GO; GO:0005355; F:glucose transporter activity; TAS.	DR	GO; GO:0005355; F:glucose transporter activity; TAS.	97	117	Extracellular (Potential).
CC	GO; GO:0005975; P:carbohydrate metabolism; TAS.	DR	GO; GO:0005975; P:carbohydrate metabolism; TAS.	118	127	4 (Potential).
CC	GO; GO:0015758; P:glucose transport; TAS.	DR	GO; GO:0015758; P:glucose transport; TAS.	128	148	Cytoplasmic (Potential).
CC	InterPro; IPR000803; Gluc_transporter.	DR	InterPro; IPR000803; Gluc_transporter.	149	156	5 (Potential).
CC	InterPro; IPR007114; MFS.	DR	InterPro; IPR007114; MFS.	157	177	6 (Potential).
CC	InterPro; IPR005828; Sub_transporter.	DR	InterPro; IPR005828; Sub_transporter.	178	182	Extracellular (Potential).
CC	InterPro; IPR005829; Sugar_transporter.	DR	InterPro; IPR005829; Sugar_transporter.	183	203	7 (Potential).
CC	InterPro; IPR003663; Sugar_transporter.	DR	InterPro; IPR003663; Sugar_transporter.	204	256	8 (Potential).
CC	Pfam; PF00083; Sugar_tr; 1.	DR	Pfam; PF00083; Sugar_tr; 1.	257	277	9 (Potential).
CC	PRINTS; PR00172; GLUCSTRNSPORT.	DR	PRINTS; PR00172; GLUCSTRNSPORT.	278	292	10 (Potential).
CC	PRINTS; PR00171; SUGRTRNSPORT.	DR	PRINTS; PR00171; SUGRTRNSPORT.	293	313	Cytoplasmic (Potential).
CC	TIGRfams; TIGR00879; SP; 1.	DR	TIGRfams; TIGR00879; SP; 1.	314	319	11 (Potential).
CC	PROSITE; PS50850; MFS; 1.	DR	PROSITE; PS50850; MFS; 1.	320	340	Cytoplasmic (Potential).
CC	PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.	DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.	341	367	Extracellular (Potential).
CC	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.	DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.	368	388	10 (Potential).
KW	Glycoprotein; Multigene family; Sugar transport; Transmembrane; Transport.	KW	Glycoprotein; Multigene family; Sugar transport; Transmembrane; Transport.	389	404	Cytoplasmic (Potential).
FT	DOMAIN	FT	DOMAIN	405	425	11 (Potential).
FT	TRANSMEM	FT	TRANSMEM	426	438	Extracellular (Potential).
FT	DOMAIN	FT	DOMAIN	439	459	12 (Potential).
FT	TRANSMEM	FT	TRANSMEM	460	477	Cytoplasmic (Potential).
FT	DOMAIN	FT	DOMAIN	478	492	Extracellular (Potential).
FT	TRANSMEM	FT	TRANSMEM	493	513	10 (Potential).
FT	DOMAIN	FT	DOMAIN	514	519	Cytoplasmic (Potential).
FT	TRANSMEM	FT	TRANSMEM	520	540	9 (Potential).
FT	DOMAIN	FT	DOMAIN	541	567	Extracellular (Potential).
FT	TRANSMEM	FT	TRANSMEM	568	588	10 (Potential).
FT	DOMAIN	FT	DOMAIN	589	604	Cytoplasmic (Potential).
FT	TRANSMEM	FT	TRANSMEM	605	625	11 (Potential).
FT	DOMAIN	FT	DOMAIN	626	638	Extracellular (Potential).
FT	TRANSMEM	FT	TRANSMEM	639	659	12 (Potential).
FT	DOMAIN	FT	DOMAIN	660	677	Cytoplasmic (Potential).
FT	SITE	FT	SITE	12	13	Dileucine internalization motif (By similarity).
FT	CARBOHYD	FT	CARBOHYD	349	349	N-linked (GlcNAc...) (By similarity).
FT	CONFLICT	FT	CONFLICT	377	377	S -> N (in Ref. 2).

DR	PROSITE; PS50850; MFS; 1.	DR	PROSITE; PS50850; MFS; 1.	100.0%	Score 2457;	DB 2;	Length 477;
DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.	DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.	100.0%	Fred. No. 8.8e-184;	Mismatches 0;	Indels 0;
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.	DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.	100.0%	Fred. No. 8.8e-184;	Mismatches 0;	Indels 0;
KW	Sugar transport; Transmembrane; Transport.	KW	Sugar transport; Transmembrane; Transport.	477 AA;	50819 MW;	OB480F94AF063316	CRC64;
SQ	SEQUENCE	SQ	SEQUENCE	477 AA;	50819 MW;	OB480F94AF063316	CRC64;
DR	Query Match	DR	Query Match	100.0%	Score 2457;	DB 2;	Length 477;
DR	Best Local Similarity	DR	Best Local Similarity	100.0%	Fred. No. 8.8e-184;	Mismatches 0;	Indels 0;
DR	Matches 477; Conservative 0;	DR	Matches 477; Conservative 0;	100.0%	Fred. No. 8.8e-184;	Mismatches 0;	Indels 0;
QY	1	QY	1	60	60		
DB	1	DB	1	60	60		
QY	61	QY	61	120	120		
DB	61	DB	61	120	120		
QY	121	QY	121	180	180		
DB	121	DB	121	180	180		
QY	181	QY	181	240	240		
DB	181	DB	181	240	240		
QY	241	QY	241	300	300		
DB	241	DB	241	300	300		
QY	301	QY	301	360	360		
DB	301	DB	301	360	360		
QY	361	QY	361	420	420		
DB	361	DB	361	420	420		
QY	421	QY	421	477	477		
DB	421	DB	421	477	477		

RESULT 2

GTRE_HUMAN STANDARD; PRT; 477 AA.

AC Q9NY64; Q9NSC4;

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Solute carrier family 2, facilitated glucose transporter, member 8 (Glucose transporter type 8) (Glucose transporter type XI).

GN Name=SLC2A8; Synonyms=GLUT8, GLUTX1;

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_taxid=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Testis;

RX MEDLINE=20283667; PubMed=10821868; DOI=10.1074/jbc.275.21.16275;

RA Dege H., Schuermann A., Bahrenberg G., Brauers A., Joost H.-G.;

RT "GLUT8, a novel member of the sugar transport facilitator family with glucose transport activity."

RL J. Biol. Chem. 275:16275-16280(2000).

RN [2]

RP SEQUENCE FROM N.A.

RA MEDLINE=20139191; PubMed=10671487; DOI=10.1074/jbc.275.7.4607;

RX Ibberson M.R., Uldry M.A., Thorens B.;

RT "GLUTX1, a novel mammalian glucose transporter expressed in the central nervous system and insulin-sensitive tissues."

RL J. Biol. Chem. 275:4607-4612(2000).

RL -1- FUNCTION: Insulin-regulated facilitative glucose transporter.

FT CONFLICT 456 457 FS -> LF (in Ref. 2).
 SQ CONFLICT 462 462 T -> I (in Ref. 2).
 SQ SEQUENCE 477 AA; 50792 MW; 0B480F94B40AEE76 CRC64;

Query Match
 Best Local Similarity 99.5%; Score 2445; DB 1; Length 477;
 Matches 475; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MTPEDPEETOPLLGPPGSGAPRRRVFLAAFAAALGFLSFGFALGYSSPAIPSLQRAAPP 60
 Db 1 MTPEDPEETOPLLGPPGSGAPRRRVFLAAFAAALGFLSFGFALGYSSPAIPSLQRAAPP 60

QY 61 APRLDDAASWFGAVVTLGAAAGVVLGGWLVDRAGRKLISLLCSVFPVAGFAVITAAQDV 120
 Db 61 APRLDDAASWFGAVVTLGAAAGVVLGGWLVDRAGRKLISLLCSVFPVAGFAVITAAQDV 120

QY 121 WMLLGGRLITGLACGVASLVAAPYIYSEIAYPAVRGLLGSVCVQLMNVVVGILLAYLAGWVLE 180
 Db 121 WMLLGGRLITGLACGVASLVAAPYIYSEIAYPAVRGLLGSVCVQLMNVVVGILLAYLAGWVLE 180

QY 181 WRWLAVLGCVPVPSMLLMLLFCMPETPRFLLTOHRRQEAAMALRFLWGSQGWEDPPIGAE 240
 Db 181 WRWLAVLGCVPVPSMLLMLLFCMPETPRFLLTOHRRQEAAMALRFLWGSQGWEDPPIGAE 240

QY 241 QSFHLALLRQPGIYKPFIIIGVSLMAFQOLSGVNAVMPFYAETIPEEAKFKDSSLASVVVGV 300
 Db 241 QSFHLALLRQPGIYKPFIIIGVSLMAFQOLSGVNAVMPFYAETIPEEAKFKDSSLASVVVGV 300

QY 301 IQVLFATAVALIMDRAGRRLLLVLSGVMVFTSFAFGAYFKLTOGGPGNSHVAISAPVS 360
 Db 301 IQVLFATAVALIMDRAGRRLLLVLSGVMVFTSFAFGAYFKLTOGGPGNSHVAISAPVS 360

QY 361 AQVSDASVGLAWLAVGSMCLFIAGFAVGVGPIPWLLMSEIFPLHVKGVATGICVLTNWL 420
 Db 361 AQVSDASVGLAWLAVGSMCLFIAGFAVGVGPIPWLLMSEIFPLHVKGVATGICVLTNWL 420

QY 421 AFLVTKFESSLMEVLRPYGAFWLASAFICFVSLVTLFCVPEPKTKLEQITAHFEGR 477
 Db 421 AFLVTKFESSLMEVLRPYGAFWLASAFICFVSLVTLFCVPEPKTKLEQITAHFEGR 477

RESULT 3
 Q6XUI2 PRELIMINARY; PRT; 478 AA.

ID O6XUI2 AC O6XUI2;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Glucose transporter 8.
 GN Name=GLUT8;
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Zhao F.-Q., Dong B., Zheng Y.;
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 RC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- SIMILARITY: Belongs to the sugar transporter family.
 DR EMBL; AY208940; AAP43920.1; -;
 DR GO; GO:0018021; C:integral to membrane; IEA.
 DR GO; GO:0005351; F:sugar porter activity; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0008643; P:carbohydrate transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub transporter.
 DR InterPro; IPR003663; Sugar_transprt.
 DR InterPro; IPR005829; Sug_transprt.
 DR Pfam; PF00083; Sugar_tr; 1.
 DR PRINTS; PR00171; SUGRTRANSPORT.

DR TIGRFAM8; TIGR00879; SP; 1.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Sugar transport; Transmembrane; Transport.
 SQ SEQUENCE 478 AA; 51416 MW; CBE71DAABCAC4CA CRC64;

Query Match
 Best Local Similarity 90.8%; Score 2230.5; DB 2; Length 478;
 Matches 425; Conservative 26; Mismatches 26; Indels 1; Gaps 1;

QY 1 MTPEDPEETOPLLGPPGSGAPRRRVFLAAFAAALGFLSFGFALGYSSPAIPSLQRAAPP 60
 Db 1 MTPEDPEETOPLLGPPGSGAPRRRVFLAAFAAALGFLSFGFALGYSSPAIPSLQRAAPP 60

QY 61 APRLDDAASWFGAVVTLGAAAGVVLGGWLVDRAGRKLISLLCSVFPVAGFAVITAAQDV 120
 Db 61 APRLDDAASWFGAVVTLGAAAGVVLGGWLVDRAGRKLISLLCSVFPVAGFAVITAAQDV 120

QY 121 WMLLGGRLITGLACGVASLVAAPYIYSEIAYPAVRGLLGSVCVQLMNVVVGILLAYLAGWVLE 180
 Db 121 WMLLGGRLITGLACGVASLVAAPYIYSEIAYPAVRGLLGSVCVQLMNVVVGILLAYLAGWVLE 180

QY 181 WRWLAVLGCVPVPSMLLMLLFCMPETPRFLLTOHRRQEAAMALRFLWGSQGWEDPPIGAE 240
 Db 181 WRWLAVLGCVPVPSMLLMLLFCMPETPRFLLTOHRRQEAAMALRFLWGSQGWEDPPIGAE 240

QY 241 QSFHLALLRQPGIYKPFIIIGVSLMAFQOLSGVNAVMPFYAETIPEEAKFKDSSLASVVVGV 299
 Db 241 HQDFHVAQLRRPGVYKPFIIIGVSLMAFQOLSGVNAVMPFYAETIPEEAKFKDSSLASVVVGV 300

QY 300 VLOVLFATAVALIMDRAGRRLLLVLSGVMVFTSFAFGAYFKLTOGGPGNSHVAISAPV 359
 Db 301 VLOVLFATAVALIMDRAGRRLLLVLSGVMVFTSFAFGAYFKLTOGGPGNSHVAISAPV 360

QY 360 SAQPVSDASVGLAWLAVGSMCLFIAGFAVGVGPIPWLLMSEIFPLHVKGVATGICVLTNWL 419
 Db 361 SMEAADTNVGLAWLAVGSMCLFIAGFAVGVGPIPWLLMSEIFPLHVKGVATGICVLTNWL 420

QY 420 MAFLVTKFESSLMEVLRPYGAFWLASAFICFVSLVTLFCVPEPKTKLEQITAHFEGR 477
 Db 421 MAFLVTKFESSLMEVLRPYGAFWLASAFICFVSLVTLFCVPEPKTKLEQITAHFEGR 478

RESULT 4
 GTR8_RAT STANDARD; PRT; 478 AA.

ID GTR8_RAT AC Q9JUZ1; Q9JMA6;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Solute carrier family 2, facilitated glucose transporter, member 8
 DE (Glucose transporter type 8) (Glucose transporter type XI).
 GN Name=Sic2a8; Synonyms=Glut8, Glutx1;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.; CHARACTERIZATION, AND MUTAGENESIS OF
 RP 12-LEU-LEU-13.
 RC TISSUE=Testis;
 RA MEDLINE=20138191; PubMed=10671487; DOI=10.1074/jbc.275.7.4607;
 RA Ibberson M.R., Udry M.A., Thorens B.;
 RT "GLUTX1, a novel mammalian glucose transporter expressed in the
 RT central nervous system and insulin-sensitive tissues.";
 RN J. Biol. Chem. 275:4607-4612(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Ishibaashi K.;
 RT "Molecular cloning of a new putative glucose transporter.";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

[3] TISSUE SPECIFICITY.
 RP MEDLINE=20283667; PubMed=10821868; DOI=10.1074/jbc.275.21.16275;
 RX Doege H., Schuermann A., Bahrenberg C., Brauers A., Joost H.-G.;
 RA "GLUT8, a novel member of the sugar transport facilitator family with
 RT glucose transport activity";
 RL J. Biol. Chem. 275:16275-16280(2000).
 CC -!- FUNCTION: Insulin-regulated facilitative glucose transporter.
 CC -!- Binds cytocholasin B in a glucose-inhibitable manner. Seems to be
 CC a dual-specific sugar transporter as it is inhibitable by
 CC fructose.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Principally
 CC intracellular. May move between intracellular vesicles and the
 CC plasma membrane. The dileucine internalization motif is critical
 CC for intracellular sequestration.
 CC -!- TISSUE SPECIFICITY: Highly expressed in adult and pubertal testis,
 CC but not prepubertal testis. Moderate expression in hypothalamus,
 CC cerebellum, brainstem, hippocampus, and adrenal gland. Lower
 CC amounts present in most other tissues.
 CC -!- SIMILARITY: Belongs to the sugar transporter family. Glucose
 CC transporter subfamily.

 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch)

 CC EMBL; AJ245985; CAB75729.1; -;
 DR EMBL; AB033418; BAA94383.1; -;
 DR InterPro; IPR008003; Gluc_transporter.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub_transporter.
 DR InterPro; IPR005829; Sug_transporter.
 DR InterPro; IPR003663; Sugar_transp.
 DR Pfam; PF00083; Sugar_tr_1.
 DR PRINTS; PR00172; GLUCTRNSP0T.
 DR PRINTS; PRO0171; SUGRTRNSP0T.
 DR TIGRFS; TIGR00879; SP; 1.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 DR GlycoProfile; MultiGene family; Sugar transport; Transmembrane;
 KW Glycoprotein; MultiGene family; Sugar transport; Transmembrane;
 KW Glycoprotein; MultiGene family; Sugar transport; Transmembrane;
 FT DOMAIN 1 25 Cytoplasmic (Potential).
 FT TRANSMEM 26 46 1 (Potential).
 FT DOMAIN 47 70 Extracellular (Potential).
 FT TRANSMEM 71 91 2 (Potential).
 FT DOMAIN 92 96 Cytoplasmic (Potential).
 FT TRANSMEM 97 117 3 (Potential).
 FT DOMAIN 118 127 Extracellular (Potential).
 FT TRANSMEM 128 148 4 (Potential).
 FT DOMAIN 149 156 Cytoplasmic (Potential).
 FT TRANSMEM 157 177 5 (Potential).
 FT DOMAIN 178 182 Extracellular (Potential).
 FT TRANSMEM 183 203 6 (Potential).
 FT DOMAIN 204 257 Cytoplasmic (Potential).
 FT TRANSMEM 258 278 7 (Potential).
 FT DOMAIN 279 293 Extracellular (Potential).
 FT TRANSMEM 294 314 8 (Potential).
 FT DOMAIN 315 320 Cytoplasmic (Potential).
 FT TRANSMEM 321 341 9 (Potential).
 FT DOMAIN 342 368 Extracellular (Potential).
 FT TRANSMEM 369 389 10 (Potential).
 FT DOMAIN 390 405 Cytoplasmic (Potential).
 FT TRANSMEM 406 426 11 (Potential).
 FT DOMAIN 427 439 Extracellular (Potential).
 FT TRANSMEM 440 460 12 (Potential).
 FT DOMAIN 461 478 Cytoplasmic (Potential).
 FT SITE 12 13 Dileucine internalization motif.
 FT CARBOHYD 350 350 N-linked (GlcNAc... (Probable).

FT	MUTAGEN	12	13	LL->AA: Changes subcellular location
FT	FT			mainly to the plasma membrane, there by
FT	CONFLICT	83	84	increasing transport activity.
FT				GG -> OGA (in Ref. 2).
SO	SEQUENCE	478 AA;	51458 MW;	95841FC1F18C9E9 CRC64;
	Query Match			88.0%; Score 2162.5; DB 1; Length 478;
	Best Local Similarity			85.4%; Pred. No. 3.5e-143;
	Matches			408; Conservative 33; Mismatches 36; Indels 1; Gaps 1;
Qy	1	MTPEDEPTEQLLPGGSRPRRVRVLAFAAALGSLGFGALGYSIPAIPSLQRAAPP	60	
		: :		
Db	1	MSPDPEPTEQLLRFSPGARPRRVRVLAFAAALGSLGFGALGYSIPAIPSLURRAPP	60	
		: :		
Qy	61	APRLDDAASAFQAVVTGGAVLGGWLVDRAGKLSLGLCSVVPVAVGFAVITAAQDV	120	
		: :		
Db	61	ALRLGDTAASWFGAVVTLGGAAAGVLLGGLDRAKLSLGLCSVVPVAVGFAVITAAQDV	120	
		: :~		
Qy	121	WMLGGRLLTGLACGVASLVAQVYIPIAVRVRVLAFAAALGSLGFGALGYSIPAIPSLQRAAPP	180	
		: :~		
Db	121	WMLGGRLLTGLACGVASLVAQVYIPIAVRVRVLAFAAALGSLGFGALGYSIPAIPSLQRAAPP	180	
		: :~		
Qy	181	WRMLAVLGCVPSSLMLLLMCPETPRFILTOHRRORAMAALRWKSGEOWEPPPIGAE	240	
		: :~		
Db	181	WRMLAVLGCVPSSLMLLLMCPETPRFILTOHRRORAMAALRWKSGEOWEPPPIGAE	240	
		: :~		
Qy	241	-QSFHLLRQGIYKPIIGVSLMAFQOLSGVNAVVFAYETIPEEAKFKDSSLASVTVG	299	
		: :~		
Db	241	HGFQLAWLRVGHVKLLIGLGMVFQOLSGVNAVVFAYETIPEEAKFKDSSLASVTVG	300	
		: :~		
Qy	300	VIQVLTAVAALIMDRAGRRLLLVSGVMVFSTAFAGYFKLTQGGPSSSHVAISAPV	359	
		: :~		
Db	301	IIQVLTAVAALIMDRAGRRLLLVSGVMVFSTAFAGYFKLTQGGPSSSHVAISAPV	360	
		: :~		
Qy	360	SAQPDASVGLAHLAVGSMCLIFAGVAVGCPFPIWLMSEIFRHHVGVATIGICVLINFL	419	
		: :~		
Db	361	SAEPADVHLGLAHLAVGSMCLIFAGVAVGCPFPIWLMSEIFRHHVGVATIGICVLINFL	420	
		: :~		
Qy	420	MAFLVTKFSSLMELVLPYGAFWLAASAFIFSVLFTLFCVPEPKGKLEQITAHFEGR	477	
		: :~		
Db	421	MAFLVTKFSSLMELVLPYGAFWLAASAFIFSVLFTLFCVPEPKGKLEQITAHFEGR	478	
		: :~		
	RESULT 5			
	GT88 MOUSE	STANDARD;		
	ID	_GT88 MOUSE	PRT;	477 AA.
	AC	Q9JIF3; Q9JJP4; Q9JUZ0,		
	DT	28-FEB-2003 (Rel. 41, Created)		
	DT	28-FEB-2003 (Rel. 41, Last sequence update)		
	DT	05-JUL-2004 (Rel. 44, Last annotated update)		
	DE	Solute carrier family 2, facilitated glucose transporter, member 8		
	DE	(Glucose transporter type 8) (Glucose transporter type X1).		
	GN	Name=Slc2a8; Synonyms=GLUT8, GlucX1;		
	OS	Mus musculus (Mouse).		
	OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
	OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
	OX	NCBI_TaxID=10090;		
	RN	[1]		
	RP	SEQUENCE FROM N.A.		
	RA	MEDLINE=20138191; PubMed=10671487; DOI=10.1074/jbc.275.7.4607;		
	RX	Ibberson M.R., Uldry M.A., Thorens B.;		
	RT	"GLUTX1, a novel mammalian glucose transporter expressed in the		
	RT	central nervous system and insulin-sensitive tissues.";		
	RL	J. Biol. Chem. 275:4607-4612(2000).		
	RN	[2]		
	RP	SEQUENCE FROM N.A.		
	RC	TISSUE=Testis;		
	RX	MEDLINE=20283667; PubMed=10821868; DOI=10.1074/jbc.275.21.16275;		
	RT	Doege H., Schuermann A., Bahrenberg C., Brauers A., Joost H.-G.;		
	RT	"GLUT8, a novel member of the sugar transport facilitator family with		
	RT	glucose transport activity";		
	RL	J. Biol. Chem. 275:16275-16280(2000).		
	RN	[3]		

SEQUENCE FROM N.A.
 RC STRAIN=129; TISSUE=Embryonic carcinoma; DOI=10.1073/pnas.97.13.7313;
 RX MEDLINE=20319023; PubMed=10860996; Cui Y., Pingsterhaus J.M.,
 RA Carayannopoulos M.O., Chi M.M.-Y., Mueckler M., Devastkar S.U., Moley K.H.,
 RA McKnight R.A., Mueckler M., Devastkar S.U., Moley K.H.,
 RT "GLUT8 is a glucose transporter responsible for insulin-stimulated
 RT glucose uptake in the blastocyst."; Proc. Natl. Acad. Sci. U.S.A. 97:7313-7318(2000).
 RL [4]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/Ola; TISSUE=Spleen; DOI=10.1006/bbrc.2001.5866;
 RX MEDLINE=21547794; PubMed=11689004; Joost H.-G., Schuermann A.,
 RA Schepers A., Doeghe H., Joost H.-G., Schuermann A.,
 RT "Mouse GLUT8: genomic organization and regulation of expression in
 RT 3T3-L1 adipocytes by glucose."; Biochem. Biophys. Res. Commun. 288:969-974(2001).
 CC !- FUNCTION: Insulin-regulated facilitative glucose transporter.
 CC Binds cytochalasin B in a glucose-inhibitable manner. Seems to be
 CC a dual-specific sugar transporter as it is inhibitable by
 CC fructose.
 CC !- SUBCELLULAR LOCATION: Integral membrane protein. Principally
 CC intracellular. May move between intracellular vesicles and the
 CC plasma membrane. The dileucine internalization motif is critical
 CC for intracellular sequestration (By similarity). Insulin induces a
 CC change in the intracellular localization and gives rise to
 CC insertion in the plasma membrane.
 CC !- TISSUE SPECIFICITY: Highest level of expression in placenta and
 CC testis. Highly expressed in adult and pubertal testis, but not
 CC prepubertal testis. Lower levels of expression in brain, liver,
 CC heart, kidney, fat and skeletal muscle.
 CC !- DEVELOPMENTAL STAGE: High expression in blastocysts.
 CC !- INDUCTION: Inhibited under glucose deprivation.
 CC !- SIMILARITY: Belongs to the sugar transporter family. Glucose
 CC transporter subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AJ245936; CAB75719.1; -
 DR EMBL; Y17802; CAB89815.1; -
 DR EMBL; AF232061; AAF78366.1; -
 DR EMBL; AJ413951; CAC88690.1; -
 DR MGD; MGI:1860103; Slc2a8.
 DR GO; GO:0005887; C:integral to plasma membrane; IDA.
 DR GO; GO:0005536; F:glucose binding; IDA.
 DR GO; GO:0005355; F:glucose transporter activity; IDA.
 DR GO; GO:0015758; P:glucose transport; IDA.
 DR GO; GO:0008286; P:insulin receptor signaling pathway; IDA.
 DR GO; GO:0001666; P:response to hypoxia; IDA.
 DR InterPro; IPR008003; Gluc_transporter.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub_transporter.
 DR InterPro; IPR005829; Sug_transporter.
 DR InterPro; IPR003663; Sugar_transpt.
 DR Pfam; PF00083; Sugar_Cr; 1.
 DR PRINTS; PR00172; GLUTRNSPORT.
 DR PRINTS; PR00171; SUGRTRNSPORT.
 DR TIGRFAMS; TIGR00879; SP; 1.
 DR PROSITE; PS50850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Glycoprotein; Multigene family; Sugar transport; Transmembrane;
 KW Transport.
 FT DOMAIN 1 25 Cytoplasmic (Potential).
 FT DOMAIN 26 46 1 (Potential).
 FT DOMAIN 47 70 Extracellular (Potential).
 FT TRANSMEM 71 91 2 (Potential).
 FT DOMAIN 92 96 Cytoplasmic (Potential).

FT	TRANSMEM	97	117	3 (Potential).
FT	DOMAIN	118	127	Extracellular (Potential).
FT	TRANSMEM	128	148	4 (Potential).
FT	DOMAIN	149	156	Cytoplasmic (Potential).
FT	TRANSMEM	157	177	5 (Potential).
FT	DOMAIN	178	182	Extracellular (Potential).
FT	TRANSMEM	183	203	6 (Potential).
FT	DOMAIN	204	257	Cytoplasmic (Potential).
FT	TRANSMEM	258	278	7 (Potential).
FT	DOMAIN	279	293	Extracellular (Potential).
FT	TRANSMEM	294	314	8 (Potential).
FT	DOMAIN	315	320	Cytoplasmic (Potential).
FT	TRANSMEM	321	341	9 (Potential).
FT	DOMAIN	342	367	Extracellular (Potential).
FT	TRANSMEM	368	388	10 (Potential).
FT	DOMAIN	389	404	Cytoplasmic (Potential).
FT	TRANSMEM	405	425	11 (Potential).
FT	DOMAIN	426	438	Extracellular (Potential).
FT	TRANSMEM	439	459	12 (Potential).
FT	DOMAIN	460	477	Cytoplasmic (Potential).
FT	SITE	12	13	Dileucine internalization motif (By similarity).
FT	CARBOHYD	350	350	N-linked (GlcNAc...) (By similarity).
FT	CONFLICT	39	39	S -> N (in Ref. 1).
FT	CONFLICT	94	94	S -> A (in Ref. 2 and 4).
FT	CONFLICT	429	429	S -> N (in Ref. 1).
SQ	SEQUENCE	477 AA;	51523 MW;	A3753PB34E452F9A CRC64;

Query Match 87.7%; Score 2156; DB 1; Length 477;
 Best Local Similarity 86.0%; Pred. No. 1e-142;
 Matches 411; Conservative 31; Mismatches 34; Indels 2; Gaps 2;

QY	1	MTPDEPETOPLLGGPPGAPGRVFLAAFAAALGPLSFGFALGYSPPAIFSLORAAPP	60
DB	1	MSPEDEPQETPLLRPEARTPRGRVFLASFAALGPLSFGFALGYSPPAIFSLRRTAPP	60
QY	61	APRLDDAASWFGAVVTLGAAAGVGLGWLVDRAGRKLSLLCSVPFVAVGAVITAAQDV	120
DB	61	ALRLGDNAASWFGAVVTLGAAAGVGLGWLVDRAGRKLSLLCSVPFVAVGAVITAAQDV	120
QY	121	WMLLGGRLTLGLACGVASIVAPVYISETAYPAVRGLGSCVQLMWWVGVILLAYLAGWVLE	180
DB	121	WMLLGGRLTLGLACGVASIVAPVYISETAYPAVRGLGSCVQLMWWVGVILLAYLAGWVLE	180
QY	181	WRWLAVLGCVPPLMLLMLCMFPEPRFLTLTHRQEAAMALRFLWGSQGWEDPPIGAE	240
DB	181	WRWLAVLGCVPPLMLLMLCMFPEPRFLTLTHRQEAAMALRFLWGSQGWEDPPIGAE	240
QY	241	QSFLHALLRQPGIYKPFIIIGVSLMAFQOLSGVNAVVFYAEITFEAKPKDSSLASVVVG	299
DB	241	HQGFQALLRRRPGIYKPLIIGIISLWVQOLSGVNAVVFYAEITFEAKPKDSSLASVVVG	300
QY	300	VLOVLTVAVALIMDRAGRRLLLVLSGVVWFSTAFAGYFKLTQGGPNSSSHVAISAPV	359
DB	301	IIQVLTVAVALIMDRAGRRLLLVLSGVVWFSTAFAGYFKLTQGGPNSSSHVAISAPV	359
QY	360	SAQPYDASVGLAWLAVGSMCLFIAGVWGPIPWLLMSEIPLHVKGVATGICVLTNWL	419
DB	360	RAEPYDVQVGLAWLAVGSMCLFIAGVWGPIPWLLMSEIPLHVKGVATGICVLTNWF	419
QY	420	MAFLVTKFSSLMVEVLRPYGAFWLASAFICFVSLFTLFCVPTKTKGLTQIHTAFEGR	477
DB	420	MAFLVTKFSSVMEVLRPYGAFWLTAAFCAISLVLTLLTVVPTKTKGLTQIHTAFEGR	477

RESULT 6
 GTR8_BOVIN
 ID GTR8_BOVIN STANDARD; PRT; 334 AA.
 AC P58354;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Solute carrier family 2, facilitated glucose transporter, member 8

Db 240 SVLFTLVVTPETKGRILEQVTAHFEGR 266
 RESULT 10
 Q8WZ05 PRELIMINARY; PRT; 248 AA.
 ID Q8WZ05
 AC Q8WZ05
 DT 01-MAR-2002 (TReMBLrel. 20, Created)
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang P.P., Zhou X.M., Jiang H.Q., Huang Y., Qin W.X., Zhao X.T.,
 RA Wan D.F., Gu J.R.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 DR EMBL; AF289587; AAU55771.1;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005215; F:glucose transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR Pfam; PF00083; Sugar tr.; 1.
 DR PRINTS; PR00172; GLUCTRNSPORT.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 KW Hypothetical protein; Transmembrane.
 SQ SEQUENCE 248 AA; 27017 MW; AA929A749BBEE558 CRC64;
 Query Match 46.0%; Score 1131; DB 2; Length 248;
 Best Local Similarity 93.0%; Pred. No. 2.6e-71;
 Matches 227; Conservative 5; Mismatches 6; Indels 6; Gaps 2;
 QY 164 MVVVGILLAYLAGWVLEWRWLVAVLGCYPPSLLMLLLMCFMPTPRFLLTQHRREQMAALR 223
 Db 1 MVVVGILLAYLAGWVLEWRWLVAVLGCYPPSLLMLLLMCFMPTPRFLLTQHRREQMAALR 60
 QY 224 FLWGSEQWEDPPIGAEQSFHLLALRROPYKPIIIGVSLMAFQQLSGVNAVMPYAEITF 283
 Db 61 FLWGSEQWEDPPIGAEQSFHLLALRROPYKPIIIGVSLMAFQQLSGVNAVMPYAEITF 120
 QY 284 EEAKFKDSSLASVAVSAPVSAQPVDSVGLAWLAVGSMCLFIAGVWGWP-IPWLLMSEIFP 402
 Db 121 EEAKFKDSSLASVAVSAPVSAQPVDSVGLAWLAVGSMCLFIAGVWGWP-IPWLLMSEIFP 180
 QY 344 QGGPGNSSHVAISAPVSAQPVDSVGLAWLAVGSMCLFIAGVWGWP-IPWLLMSEIFP 402
 Db 181 QGGPGNSSHVAISAPVSAQPVDSVGLAWLAVGSMCLFIAGVWGWP-IPWLLMSEIFP 235
 QY 403 LHVX 406
 Db 236 LHLQ 239
 RESULT 11
 ID Q8SPS1 PRELIMINARY; PRT; 246 AA.
 AC Q8SPS1
 DT 01-JUN-2002 (TReMBLrel. 21, Created)
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Glucose transporter 8 (Fragment).
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=CZECH II; TISSUE=Mammary tumor;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosa S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalaška U., Smalil D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Matra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CZECH II; TISSUE=Mammary tumor;
 RA Strausberg R.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- SIMILARITY: Belongs to the sugar transporter family.
 DR EMBL; BC021758; AAH21758.1;
 DR MGD; MGI:1860103; Slc2a8.
 DR GO; GO:0005887; C:integral to plasma membrane; IDA.
 DR GO; GO:0005336; F:glucose binding; IDA.
 DR GO; GO:0005335; F:glucose transporter activity; IDA.
 DR GO; GO:0006006; P:glucose metabolism; TAS.
 DR GO; GO:0015758; P:glucose transport; IDA.
 DR GO; GO:0008286; P:insulin receptor signaling pathway; IDA.
 DR GO; GO:0001666; P:response to hypoxia; IDA.
 DR Pfam; PF00083; Sugar tr.; 1.
 DR PRINTS; PR00171; SUGRTRNSPORT.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 KW Sugar transport; Transmembrane; Transport.
 FT NON_TER 1
 SQ SEQUENCE 266 AA; 29132 MW; 1B3031928863B0FC CRC64;
 Query Match 47.5%; Score 1167; DB 2; Length 266;
 Best Local Similarity 83.1%; Pred. No. 8.6e-74;
 Matches 222; Conservative 20; Mismatches 23; Indels 2; Gaps 2;
 QY 212 QHRREQMAALRFLWGSEQWEDPPIGAE-QSFHLLALRROPYKPIIIGVSLMAFQQLS 270
 Db 1 QHQIQEAMAAALRFLWGSEGWEEPPVGAHQGFQALALRRPGLYKPLIIGISLWVFOQLS 60
 QY 271 GVNAVMPYAEITFEAKFKDSSLASVAVSAPVSAQPVDSVGLAWLAVGSMCLFIAGVWGWP 330
 Db 61 GVNAVMPYANSIFEEAKFKDSSLASVAVSAPVSAQPVDSVGLAWLAVGSMCLFIAGVWV 120
 QY 331 FSTSAFGAYKPLIQGGPGNSSHVAISAPVSAQPVDSVGLAWLAVGSMCLFIAGVWGWP 390
 Db 121 FMSAEGTYFKLTKQSLPNSSSHVGL-VPIAAEPPVDVQVGLAWLAVGSMCLFIAGVWGWP 179
 QY 391 PIPWLLMSEIFPPLHVKGVAIGICVLTNWLMFAFLVTKFSSLMELVLPYGAFLMASAFCIF 450
 Db 180 PIPWLLMSEIFPPLHVKGVAIGICVLTNWLMFAFLVTKFSSVMELRYPYGAFLWTAFCAL 239
 QY 451 SVLFTLVVTPETKGRILEQVTAHFEGR 477

RC TISSUE=Placenta; DOI=10.1016/j.placenta.2003.08.012;
 RX PubMed=15013641; Regnault T.R.H., Hay W.W. Jr.,
 RA "Characterization of glucose transporter 8 (GLUT8) in the ovine
 RT placenta of normal and growth restricted fetuses.";
 RL Placenta 25:70-77(2004).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -!- SIMILARITY: Belongs to the sugar transporter family.
 DR EMBL; AF495799; AA018513.1;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005351; F:sugar porter activity; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0008643; P:carbohydrate transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub_transporter.
 DR InterPro; IPR003663; Sugar transp.
 DR InterPro; IPR005829; Sug_transporter.
 DR Pfam; PF00083; Sugar tr_1
 DR PRINTS; PR00171; SUGTRNSPORT.
 DR PROSITE; PS50850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 KW Sugar transport; Transmembrane; Transport.
 FT NON_TER 246 246
 SQ SEQUENCE 246 AA; 26537 MW; E8F3871946CD0723 CRC64;

Query Match 45.8%; Score 1125.5; DB 2; Length 246;
 Best Local Similarity 88.2%; Pred. No. 6.4e-71;
 Matches 217; Conservative 11; Mismatches 17; Indels 1; Gaps 1;

QY 219 MAALRLWSEQGWEDPPTGAE-QSFHALLRPPGKYKFFIIGVSLMARFQQLSGVNAVMP 277
 Db 1 MAAMQPLWGSAPGWEPEPTGAHQGHVAVLQRCPCGYKFFIIGISLMAFQQLSGVNAVMP 60

QY 278 YAETIFEAKFKDSSLASVVGVVQVLTAVAAALMDRAGRLLLVLSGVVWVFTSARG 337
 Db 61 YAETIFEAKFKDSSLASVVGVVQVLTAVAAALMDRAGRLLLVLSGVVWVFTSARG 120

QY 338 AYFKLTOGGPGNSSHVAISAPVSAQPDASVGLAWLAVGSMCLFIAGFAVGVGPIPWLLM 397
 Db 121 AYFKLTEGGPSNSSHVDLPAPVSMEPANTNVGLAWLAVGSMCLFIAGFAVGVGPIPWLLM 180

QY 398 SEIFPLHVKGATVCIVLNWLMFLVTFKFSLMELVLPYGAFWLASAFCIFSVLFTFLF 457
 Db 181 SEIFPLHVKGATVGVCLTNWFMAFLVTKFSSLMELVLPYGAFWLASAFCIFGVLFLLA 240

QY 458 CVPETK 463
 Db 241 CVPETK 246

RESULT 12
 GTR6_HUMAN
 ID_GTR6_HUMAN STANDARD; PRT; 507 AA.
 AC Q9UGQ3;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Solute carrier family 2, facilitated glucose transporter, member 6
 DE (Glucose transporter type 6) (Glucose transporter type 9).
 GN Names=SLC2A6; Synonyms=GLUT9;
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leukocyte;
 RX MEDLINE=20427701; PubMed=10970791; DOI=10.1042/0264-6021:3500771;
 RA Doeghe H., Bocianski A., Joost H.-G., Schuermann A.;
 RT "Activity and genomic organization of human glucose transporter 9
 (GLUT9), a novel member of the family of sugar-transport facilitators
 predominantly expressed in brain and leucocytes.";
 RL Biochem. J. 350:771-776(2000).

RN SEQUENCE FROM N.A.
 RP TISSUE=Lymphoid;
 RA Young J.M., Woodward K.J., Aziz S., Burley M., Kwiatkowski D.J.,
 RA Povey S.,
 RT "Cloning of a sugar transporter gene, a G-beta subunit like gene and
 RT three novel genes in human chromosome 9q34.";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulláhy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Hellon E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: Facilitative glucose transporter; binds cytochalasin B
 CC with low affinity.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. The dileucine
 CC internalization motif is critical for intracellular sequestration
 CC (By similarity).
 CC -!- TISSUE SPECIFICITY: Highly expressed in brain, spleen and
 CC peripheral blood leukocytes.
 CC -!- SIMILARITY: belongs to the sugar transporter family. Glucose
 CC transporter subfamily.
 CC -!- CAUTION: Has been described as GLUT9 in literature, but this gene
 CC name has already been used for SLC2A9.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Y17803; CAB96996.1; --
 DR EMBL; AJ011372; CAB66155.1; --
 DR EMBL; BC013740; AAHL3740.1; --
 DR Genbank; HGNC:11011; SLC2A6.
 DR H-InvDB; HIX0008509; --
 DR MIM; 606813; --
 DR InterPro; IPR000803; Gluc_transporter.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub_transporter.
 DR InterPro; IPR005829; Sug_transporter.
 DR InterPro; IPR003663; Sugar_transp.
 DR Pfam; PF00083; Sugar tr_1
 DR PRINTS; PR00172; GLUCTRNSPORT.
 DR PRINTS; PR00171; SUGTRNSPORT.
 DR TIGRFAMs; TIGR00879; SP; 1.
 DR PROSITE; PS50850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Glycoprotein; Multigene family; Sugar transport; Transmembrane;
 KW Transport.
 FT DOMAIN 1 37 Cytoplasmic (Potential).

FT TRANSMEM 38 58 1 (Potential).
 FT DOMAIN 59 81 Extracellular (Potential).
 FT TRANSMEM 82 102 2 (Potential).
 FT DOMAIN 103 111 Cytoplasmic (Potential).
 FT TRANSMEM 112 132 3 (Potential).
 FT DOMAIN 133 140 Extracellular (Potential).
 FT TRANSMEM 141 161 4 (Potential).
 FT DOMAIN 162 168 Cytoplasmic (Potential).
 FT TRANSMEM 169 189 5 (Potential).
 FT DOMAIN 190 194 Extracellular (Potential).
 FT TRANSMEM 195 215 6 (Potential).
 FT DOMAIN 216 289 Cytoplasmic (Potential).
 FT TRANSMEM 290 310 7 (Potential).
 FT DOMAIN 311 314 Extracellular (Potential).
 FT TRANSMEM 315 335 8 (Potential).
 FT DOMAIN 336 339 Cytoplasmic (Potential).
 FT TRANSMEM 340 360 9 (Potential).
 FT DOMAIN 361 395 Extracellular (Potential).
 FT TRANSMEM 396 416 10 (Potential).
 FT DOMAIN 417 435 Cytoplasmic (Potential).
 FT TRANSMEM 436 456 11 (Potential).
 FT DOMAIN 457 462 Extracellular (Potential).
 FT TRANSMEM 463 483 12 (Potential).
 FT DOMAIN 484 507 Dileucine internalization motif (By similarity)
 FT SITE 5 6 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 370 370 M -> T (in Ref. 2).
 FT CONFLICT 500 500
 FT SEQUENCE 507 AA; 54569 MW; 1FC6E2EB11588460 CRC64;

Query Match 38.6%; Score 948; DB 1; Length 507;
 Best Local Similarity 42.7%; Pced. No. 3.4e-58;
 Matches 216; Conservative 80; Mismatches 166; Indels 44; Gaps 11;

QY 10 QLLG-----PPGSAPGR-----RVFLAAFAAALGPLSPFCALGYSRPA 50
 Db 3 ELLGAEQDYDTPPEKPPSPGDRARVGTQKRVFLATFAVLGNFSGYALVYTSVP 62
 QY 51 IFSLQRAAPRFLDDAASWFGAVVTGLGAAGGVLGGWLVDRAGKLSLLCSVFFVAG 110
 Db 63 IFLERSLDDHLTKSQASWFSVFTLGAAGGSLMFLNLLGRKLSIMFSAVPSAAG 122
 QY 111 FAVITAAQDVMMLGRLTGLACGVASIVAPVYISEIAYPAVRGLGSCVQVMVVVGL 170
 Db 123 YALMAGAHGLWMLLGRUTGTGAGGLTAAICIPVYVSEIAPGVRGALGATPOLMAVFGSL 182
 QY 171 LAYLAGWLEWRWLVGCVPSMLLMCFMPETPRELLTQHRQERMAALRFLWVSEQ 230
 Db 183 SLYALGLLPLWRWLVAVAGEAPVLMILLISFNPSPRFLSRGRDEEARLAWLRGTDV 242
 QY 231 G--WEDPPIG---AEQSFLH--ALLRQGIYKPFIIIGVSLMAFOOLSGVNAVMFYAETIF 283
 Db 243 DVHWEFEQLDNRVQRSSVWAERAEVHCVRPITVALLMRLLOQTGHTPLFVYLQSF 302
 QY 284 EBAKE---KDSLSAVVGVVQVLFVFAALIMDRGRRLLLVLSGVVMVFTSAFAGY 339
 Db 303 DSTAVLLPKDD---AAIVGAVRLLSVLIATLMDLAGRKVLLFVSAIMFAANLTLGLY 359
 QY 340 FKLTOGGFGNSHVAISAPVS-----AQPVDASVG-LAWLAVGSMCLFIAGFVAGWGP 392
 Db 360 IHF---GPRPLSPNNTAGLESWGDLAQLAAPAGYLLTFLPMTLFLMFGYAVGWGPI 416
 QY 393 PWWLMSFPLHVKGVATGICVLTWLMMAFLVTKFSSIMEVLRPYGAFWLASAPCFISV 452
 Db 417 TWLLMSEVLPLEARGVAGSLCVLAWLTAFLVTKSFLPVVSTFGLQVFFPFAICLVSL 476
 QY 453 LFTLFCVETPKGKTLLEQITAHPE--GR 477
 Db 477 VFTGCCVETPKGRSLEQIESFRMGR 502

RESULT 13
 QBSTN2

ID QBSTN2 PRELIMINARY; PRT; 497 AA.
 AC QBSTN2;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Mus musculus NOD-derived CD11c +ve dendritic cells cDNA, RIKEN full-length enriched library, clone:F63103L12 product:similar to SOLUTE CARRIER FAMILY 2, FACILITATED GLUCOSE TRANSPORTER, MEMBER 6 (GLUCOSE TRANSPORTER TYPE 6) (GLUCOSE TRANSPORTER TYPE 9).
 DE TRANSPORTER TYPE 6)
 GN Name=S1c2a6;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NOD;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NOD;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/350555500;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NOD;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NOD;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Kanno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NOD;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P., Kanno H., Akiyama J., Nishi K., Kitsuma T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A., Yanamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M., Onoda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NOD;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P., Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayato K., Hiramoto K., Hiraoka T., Hirozane T., Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T., Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S., Kurihara C., Matsuuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y., Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akaira S., Takeda Y., Tanaka T., Tomaru A., Toyota T., Yaeunishi A., Muramatsu M., Hayashizaki Y.;

Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -!- SIMILARITY: Belongs to the sugar transporter family.
 DR EMBL; AK089246; BAC40811.1; -.
 DR MGD; MGI:2443286; SLC2a6.
 DR GO; GO:0016020; C-membrane; ISS.
 DR GO; GO:0005355; F:glucose transporter activity; ISS.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub transporter.
 DR InterPro; IPR003663; Sugar transp.
 DR InterPro; IPR005829; Sug transporter.
 DR Pfam; PF00083; Sugar_tr; 1.
 DR PRINTS; PRO0171; SUGTRNSPORT.
 DR TIGRFAMs; TIGR00879; SP; 1.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Sugar transport; Transmembrane; Transport.
 SQ SEQUENCE 497 AA; 54445 MW; 8E14D1DCD4B244E3 CRC64;

Query Match 38.5%; Score 946; DB 2; Length 497;
 Best Local Similarity 43.6%; Pred. No. 4.6e-58;
 Matches 214; Conservative 77; Mismatches 168; Indels 32; Gaps 9;

QY 3 PEDPEETQPLGGPPG-----GSAPRRRVFLAFAAALGFLSFGFALGYSSPAIPSLQRAA 58
 Db 17 PEVP-----ATPGRERAGALKNRRVFLATFAAVLGNFSGYALVYTSVPIPELKLSS 69
 QY 59 PPAFLDDAAASWFGAVVILGAAAGVIGLWVDRAGRKLKLLSLLCSVFPVAGFAVITAAQ 118
 Db 70 DPALHLDKIQASWFGSVFTLGAAGLSAMLLNLLGRKLSIMFSAPSAIGYAIMAGAR 129
 QY 119 DVYMLLGGRLTLGLACCVASLAVPVVYISIAVPAVRGLGSCVQLMNVVGGILLAVLAGWV 178
 Db 130 GYMLLGLRMLTGFAGGLTAACIPVYVSEIAPDVYRGALGATPQLMNVFSGLSLVALGLL 189
 QY 179 LEWRMLAVLGCVPVPSLMLLLMCFMPTFRLLTQHRREQEAMALRFLWG-SEQGWEDPPI 237
 Db 190 LPWRMLAVAGEPVLIMILLSPMNSPRFLLSKSRDEALQALTLWRADSEVHVEFEQI 249
 QY 238 G---AEQSHL--ALLRQGIYKFFIIGYSLMAFQOLSGVNAVVFYAEIPEAK-FKDS 291
 Db 250 QDNVRRQSRVSWAEAREPRVYRVLIIAVMRFLOQLTGITPILVYLQIFDNTSVVLPFS 309
 QY 292 SLASVVVQIVLFTAAALIMDRAGRLLLVLSGVVWVVFSTSAFGAYFK-----LTQGG 346
 Db 310 QQDAALVGVAVRLLSVLIIAATVMDLAGRKLIVYSASVMPAANLTLGLYVQVFRPLT--- 366
 QY 347 PGNSSHVAISAPVSAQPVDSVGLAWLAVGSMCLFIAGFVAGVGPVPLMSEIIFPLHYK 406
 Db 367 PNSTVEIVLGD-----AFNYLTLIPLLATMFLFMGYMGWGPITWLLMSEVLPRLAR 420
 QY 407 GVATGICVLTWMLMAFLVTKPSSLMNEVLRPGAFWLASAFCLFSLVFLFCVPETKGT 466
 Db 421 GVASGLCVLVSMLTAFVNLVYLLAVNAFGLQVPPFFFAICLLSLLTGGCCVPETRGRS 480
 QY 467 LEQITAHFEGR 477
 Db 481 LEQIEAFFHTR 491

OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.,
 RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative";
 RL Dev. Dyn. 225:384-391(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Buetow K.H., Wang J., Hsieh F.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ushin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalios D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RL and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Klein S., Gerhard D.S.;
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 DR EMBL; BC081076; AAH81076.1; -.
 DR InterPro; IPR000803; Gluc_transporter.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub transporter.
 DR InterPro; IPR003663; Sugar transp.
 DR InterPro; IPR005829; Sug transporter.
 DR Pfam; PF00083; Sugar_tr; 1.
 DR PRINTS; PRO0172; GLUCTRNSPORT.
 DR PRINTS; PRO0171; SUGTRNSPORT.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Sugar transport; Transmembrane; Transport.
 SQ SEQUENCE 465 AA; 51045 MW; AE92305C299BC562 CRC64;

Query Match 37.8%; Score 928; DB 2; Length 465;
 Best Local Similarity 41.9%; Pred. No. 7.8e-57;
 Matches 198; Conservative 90; Mismatches 152; Indels 32; Gaps 6;

QY 26 VFLAAFAAALGFLSFGFALGYSSPAIPSLQRAAPPAPRLDDAA-----ASWFGAVVTL 78
 Db 1 MFLAAFSAVLGNFTFGYALVYTSVPIPALEK-----DDQGLHINASEIISWFGAVFAL 52
 QY 79 GAAAGVGGWLVDRAGRKLKLLSLLCSVFPVAGFAVITAAQDVYMLLGGRLTLGLACCVAS 138
 Db 53 GACAGGISMFLNDRLGRKLSIMFSAPVPSLGLMGLMGSQAHSMLLGLRILTFAGGWTSS 112
 QY 139 LVAVYVISEIAYPAVRGLGSCVQLMNVVGGILLAVLAGWVLEWRMLAVLGCVPVPSLMLLL 198
 Db 113 SSIPVYIISHSVGRGLGACVQPMVAVCGSLVYALGULLPWRMLAIGEVVYVYVTLMLL 172
 QY 199 MCFMPTFRLLTQHRREQEAMALRFLWGSEQGWEDPPIGASQSFHALLR----- 249

RESULT 14
 Q66J34
 ID PRELIMINARY; PRT; 465 AA.
 AC Q66J34;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE MGC82056 protein.
 GN Name=MGC82056;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;

```

Db 173 LCFMPSDFRLIAKGDKALALWLRGANTDYQ---GEYERIKSNILKSSSTLWTE 228
Qy 250 --OPGYKPFIVGSLMFCQQLSNGVNAVMPYAEIPEEAK-FKDSLSLAVVVVGIQVLF 306
Db 229 LSPYKPIKILIAVFNRFQQLSGVSPILILETIPNRTKVLIRGGYDAALVGVVRLLSV 288
Qy 307 AVAALIMDRAGRLILVLSGVVVFSTSAFAYFKLTQGGPNSHVAISAPVSAQPVDA 366
Db 289 IISASVMDKAGRKILLYSTLMSVSSLSMGLYVHTVDINHNSNTRTWISISSAEPSEP 348
Qy 367 SYGLAWLAVGSMCLFIAGVAGVGPVPIVLLMSEIIFPLHVKGVATGICVLTNMLMAFLVTK 426
Db 349 VNVYIQLLILICIMLYIIGVAFNGVPIVLLMSEIIFPLHVKGVATGICVLTNMLMAFLVTK 408
Qy 427 EFSLSMEVLRPFGAFWLASAFICFVFLFCVPEPKTKTLQIHTAHFE-GR 477
Db 409 AFIPVVNTLSLQTPFYFFTAACAASIMFTYFFVPEPKTKTLQIHTAHFE-GR 460

```

RESULT 15

```

QBNC22
ID QBNC22 PRELIMINARY; PRT; 445 AA.
AC QBNC22;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein FLJ90355.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
SEQUENCE FROM N.A.
RA Isoyagi T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,
RA Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RL Submitted (MAR-2002) to the EMBL/Genbank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005355; F:glucose transporter activity; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; F:carbohydrate transport; IEA.
DR InterPro; IPR000803; Gluc_transporter.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005028; Sub_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005029; Sug_transporter.
DR Pfam; PF00083; Sugar_cr; 1.
DR PRINTS; PR00172; GLUCRNSPORT.
DR PRINTS; PR00171; SUGRTRNSPORT.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transporter; Transmembrane; Transporter.
SQ SEQUENCE 445 AA; 48040 MW; BF37DC0C313A32CE CRC64;

```

```

Query Match 36.4%; Score 894; DB 2; Length 445;
Best Local Similarity 40.1%; Pred. No. 1.8e-54;
Matches 200; Conservative 73; Mismatches 134; Indels 92; Gaps 9;
Qy 10 QPLIG-----PPGGSAPRGR-----RVFLAAFAAALGPLSGFALGYSSPA 50
Db 3 EPLLGAEGPDYDTFPEKPPPPGDRARVGTLQNKVFLATEFAAVLGNFSGYALVYTSV 62
Qy 51 IPSLQRAAPRDLDAASWFGAVVTIIGAAGVGLGWLVDRAGRKLSLLLCVSPFFVAG 110
Db 63 IPALERSLDPDLHLTKSQASWFGVFTLGAAGGUSAMILNDLGRKUSIMFSAVPSAAG 122

```

```

Qy 111 FAVITAAQDVWMLLGRLLTGLACGVASLVAAPVYISETAYPAVRGLGSCVQLMNVVGLI 170
Db 123 YALMAGAHLWMLLGRLLTGFAGGLTAACIPVYVSEIAPPVGRGALGATPQLMAVFGSL 182
Qy 171 LAYLAGWLEWRWLAFLVGCVPSPSMLLMLCFMPEPFRLLTQHRROEAMAALRFLWGSQ 230
Db 193 SLIYALGLLLPWRWLAAGEAPVLIWMLLMLLSPFNSPFRLLSRGRDEEALRALAWLRGTDV 242
Qy 231 G--WEDPPIG---AEQSFHL--ALLRQPIYKPFIIYVSLMAFQQLSGVNAVMPYAEIIF 283
Db 243 DVHWEFEIQDNRVROSSRVSWAEARAPHVCRPIVALLMRLQLQTTGTTPLVYLQSF 302
Qy 284 EAKF---KDSLSLAVVVVGIQVLFVAVALIMDRAGRLLLVLSGVVMPFSTSAFGAY 339
Db 303 DSTAVLLPPKOD---AAIVGAVRLLSVLIIAALTMDLAGRKVL----- 341
Qy 340 FKLITQGGPNSHVAISAPVSAQPVDAVGLAWLAVGSMCLFIAGFAYGWPPIVLLMSE 399
Db 342 -----LFGVGYAVGWGPIVLLMSE 361
Qy 400 IFFLHVKGVATGICVLTNMLMAFLVTKFSSLMVLRPYGAFWLASAFICFVSLVFLFCV 459
Db 362 VLPLRARGVASGLCVLASWLTAFVLTKSFLPVPVSTFGLQVDFEFAAICLVSLVFTGCCV 421
Qy 460 PETKGTLEQIHTAHFE-GR 477
Db 422 PETKGRSLEQIESPFRTRGR 440

```

Search completed: February 24, 2005, 05:52:57
Job time : 183 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
February 25, 2005, 16:32:02 ; Search time 5101 Seconds
(without alignments)
4531.099 Million cell updates/sec

Title: US-09-886-954A-1
Perfect score: 2457
Sequence: 1 MTPEDPEFQPLLPGGSA.....CVPETKGTLEQITAHFEGR 477

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=slp
-Q/cgn2_1/USPTO_spool_P/US09886954/runat_23022005_154242_13208/app_query_faasta_1.647
-DB=GenEmbl -QFWT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdd -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09886954 @CGN 1.1 5600 @runat_23022005_154242_13208 -ICPU=3
-NO MWAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_cv:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: sp_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Table with 5 columns: Result No., Score, Query Match, Length, Description. Contains 4 rows of search results.

Table with 5 columns: Line number, Score, Query Match, Length, Description. Contains 45 rows of search results.

ALIGNMENTS

RESULT 1 AX179740 2080 bp DNA linear PAT 06-AUG-2001
AX179740
LOCUS Sequence 28 from Patent WO0146258.
DEFINITION AX179740
ACCESSION AX179740
VERSION AX179740.1 GI:15132104
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
Baughn, M.R., Burford, N., Au-Young, J., Lu, D.A., Yang, J., Reddy, R.,
Lal, P., Hillman, J.L., Azimzal, Y., Yue, H., Nguyen, D.B., Yao, M.G.,
Gandhi, A.R., Tang, Y.T. and Khan, F.A.
Transposers and ion channels
Patent: WO 0146258-A 28 28-JUN-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
i. .2080
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="inctype ID No: 1416107CBI"

Alignment Scores: 7.58e-167 Length: 2080
Pred. No.: 2080

Score: 2457.00 Matches: 477
 Percent Similarity: 100.00% Conservativity: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-886-954A-1 (1-477) x AX179740 (1-2080)

QY 1 MetThrProGluAspProGluGlnThrGlnProLeuLeuGlyProProGlyGlySerAla 20
 |||
 Db 46 ATGACGCCCGGAGACCAGGAAACCCAGCGCTTCGTGGCCCTCCCTGGCGGAGCGG 105
 |||
 QY 21 ProArgGlyArgValPheLeuAlaPheAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
 |||
 Db 106 CCCCGCGCGCGCGCTTCCTCCCGCCCTTCGCGCGCTTCGCGCGCTTCGCGCGCTTC 165
 |||
 QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 |||
 Db 166 GCGTTTCGCGCTCGGCTACAGCTCCCGGCAATCCCTAGCCTGCAGCGCGCGCGCCCG 225
 |||
 QY 61 AlaProArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrIleuGlyAla 80
 |||
 Db 226 GCGCGCGCTGAGAGCGCGCGCTTCCTCGGCTTCGCGCGCTTCGCGCGCTTCGCGCG 285
 |||
 QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgGlyLeuSerLeu 100
 |||
 Db 286 GCGCGCGCGGAGTGTCTGGCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCG 345
 |||
 QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 |||
 Db 346 TTGCTGTGCTTCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCG 405
 |||
 QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 |||
 Db 406 TGGATGCTGCGCGCGCGCGCTTCCTCAGCGCGCTTCGCGCGCTTCGCGCGCTTCGCG 465
 |||
 QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuLeuSerCys 160
 |||
 Db 466 GCGCGCGCTTACATATCCGAAATCCGCTTACCAGCAGTCCGCGCGCTTCGCGCGCT 525
 |||
 QY 161 ValGlnLeuMetValValValGlyIleLeuLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
 |||
 Db 526 GTGACGTAATGTCGTCGCGCATCTCTCGCGCTTCGCGCGCTTCGCGCGCTTCGCG 585
 |||
 QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 |||
 Db 586 TGGCGCTGGTGGTGTGCTGGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGG 645
 |||
 QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
 |||
 Db 646 TTCATGCCGAGACCCCGCTTCCTGCTGACTCAGCAGCGCGCGCGCGCGCGCGCGCG 705
 |||
 QY 221 AlaLeuArgPheLeuTrpGlySerGlnGlyTrpGluAspProIleGlyAlaGlu 240
 |||
 Db 706 GCGCTGGTTCCTGTGGGCTCGGAGCGGCTGGGAGACCCCGCGCGCGCGCGCGCGCG 765
 |||
 QY 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIleValPheIleIleGly 260
 |||
 Db 766 CAGAGCTTTCACCTGGCGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 825
 |||
 QY 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
 |||
 Db 826 GTCTCCCTGATGGCTTCAGCAGCTGTCGCGCGCTCAACCGCTCATGTTCTATGCGAG 885
 |||
 QY 281 ThrIlePheGluGluAlaLysPheIysAspSerLeuAlaSerValValValGlyVal 300
 |||
 Db 886 ACCATCTTTGAAGAGGCCAAAGTTCAGAGCAGCAGCTTCGCGCGCTTCGCGCGCTTC 945
 |||
 QY 301 IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgLeu 320
 |||
 Db 946 ATCCAGGTGCTTTCACAGCTGTCGCGCGCTTCATCATGAGCAGCGCGCGCGCGCG 1005
 |||
 QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340

1006 CTCCTGGCTTTCAGGTGTGTCATGATGTTTCAGCAGGAGTGCCTTCGGCGCTACTTTC 1065
 |||
 QY 341 LysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer 360
 |||
 Db 1066 AAGCTGACCCAGGCTGGCCCTGCACTCTCCGACGCTGGCCATCTTCGGCGCTGTCTCT 1125
 |||
 QY 361 AlaGlnProValAspAlaSerValGlyLeuAlaTlpLeuAlaValGlySerMetCysLeu 380
 |||
 Db 1126 GCACAGCCCTGTGATCCAGCGTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1185
 |||
 QY 381 PheIleAlaGlyPheAlaValGlyTrpGlyTrpGlyProIleProTrpLeuLeuMetSerGluIle 400
 |||
 Db 1186 TTCATCGCGCGCTTCGCGCTGGCGTGGCGCCCATCCCTGGCTCCCTCATGTCAGAGATC 1245
 |||
 QY 401 PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
 |||
 Db 1246 TTCCTCTGCAATCAAGGGCGTGGCGCAGGCAATCTGGCTCCCTCACCACTGGCTCATG 1305
 |||
 QY 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
 |||
 Db 1306 GCCTTTCGTGACCAAGGAGTTCAGCAGCTCATGAGGCTTCATGAGGCTTCATGAGGC 1365
 |||
 QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
 |||
 Db 1366 TTCCTGGCTCCCGCTTCGCAATCTGCAATCTGCAATCTGCAATCTGCAATCTGCAAT 1425
 |||
 QY 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 |||
 Db 1426 GAAACTAAGAAAGACTCTGGAACTGGAACTGGAACTGGAACTGGAACTGGAACTGG 1476
 |||

RESULT 2
 BC019043 2145 bp mRNA linear PRI 30-JUN-2004
 LOCUS BC019043
 DEFINITION Homo sapiens solute carrier family 2, (facilitated glucose transporter) member 8, mRNA (cDNA clone MGC:20634 IMAGE:4641145), complete cds.

ACCESSION BC019043
 VERSION BC019043.2 GI:40226554
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 2145)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klusner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Alexchul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S., Carninci, P., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Green, E.D., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Schmutz, J., Myers, R.M., Dickson, C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Butlerfield, Y.S., Krzywicki, M.I., Skalska, U., Smallos, D.E., Smerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE
 JOURNAL PUBLISHED
 12477932
 REFERENCE 2 (bases 1 to 2145)
 Strausberg, R.
 Direct Submission
 Submitted (07-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 On Dec 19, 2003 this sequence version replaced gi:17512129.
 Contact: MGC help desk
 Email: sgaps-@email.nih.gov
 Tissue Procurement: ATCC/DCTP/DTP
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nhgri.nih.gov
 Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghghi, P.,
 Hansan, N., Ho, S.-L., Karins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantropop, S., Thomas, P.J., Touchman, J.W.,
 Tsutjeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL.Plate: 30 Row: n Column: 3
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 21361448.

FEATURES
 Location/Qualifiers
 1..2145
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:20634 IMAGE:4641145"
 /tissue_type="Skin, melanotic melanoma."
 /clone_lib="NIH_MGC_20"
 /lab_host="DH10B-R"
 /notes="Vector: pOTB7"
 1..2145
 /gene="SLC2A8"
 /note="synonyms: GLUTX1, GLUT8"
 /db_xref="LocusID:29988"
 /db_xref="MIM:605245"
 46..1479
 /gene="SLC2A8"
 /codon_start=1
 /product="solute carrier family 2, (facilitated glucose
 transporter) member 8"
 /protein_id="AAH19043.1"
 /db_xref="GI:17512130"
 /db_xref="LocusID:29988"
 /db_xref="MIM:605245"
 /translation="MTPEDPEETPLLPPGGSAPRRRVLAAFAALGPLSFGFAL
 GYSSPAIPSLQRAAPPRLDDAAASWFGAVVTLGAAGVGGWLVDRGRKLSLL
 CSVPVAGFAVITAAQDVMWLLGGRLITGLACGASVAVPVYISIAVPAVRLGLGC
 VOLMVGGLLALAGWLEWRWLVGCVPSLMLLLMFMPEPTFRLLTTHRRLOEA
 MAALRWLGSQGWEDPPIQAFOSFHLLRQPLKPIIGVSLMAFQGLGVNAV
 FYAFTIFEEKFKESLAVSVVQLVFTVAALLMDRAGRLLLVLSGVYVWVST
 AFGAFKIQGGPNSHVAISAPVSAQVDSVGLWLAAGVSMCLFAGFAVWGPI
 PWLLMSEIFPLHVGVATGICVLTWNMAFLVTKGFSSLMELVIRPYGAPFWLASAF
 SVLFTLFCVPEPKTKGLEITAFHEGR"

Alignment Scores:
 Pred. No.: 7,86e-167 Length: 2145
 Score: 2457.00 Matches: 477
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

ORIGIN
 1 MetProGluAspProGluThrGlnProLeuLeuGlyProGlyGlySerAla 20

Db	46	ATGACGCCCGAGGACCCAGAGAAACCCAGCGCTTCTGGGGCTCTCTGGCGGACGCGG 105
Qy	21	ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
Db	106	CCCCCGCGCGCGCGCTTCTCGCGCGCTTTCCTGGCGCTTTCCTGGCGCGCGCGCTT 165
Qy	41	GlyPheAlaLeuGlyTyrSerProAlaLeuProSerLeuGlnArgAlaAlaProPro 60
Db	166	GGCTTCGCGCTTGGCTACAGCTCCCGCGCACTCCCTAGCTGCGCGCGCGCGCGCG 225
Qy	61	AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
Db	226	GCCTCGCGCTGGAGAGCGCGCGCTCTCTGGTTCGGGGCTGCTGTCGACCTGGGTGC 285
Qy	81	AlaAlaGlyGlyValLeuGlyTrpLeuValAspArgAlaGlyArgIysLeuSerLeu 100
Db	286	CGCGCGCGGGAGTGTGGCGCGCTGGCTGGTGGACCGCGCGCGCGCGCGCGCGCG 345
Qy	101	LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
Db	346	TTGCTGTGCTCGTGGCGCTTCTGGCGCGCTTTCCTGCTCATCACCGCGCGCGCGCG 405
Qy	121	TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
Db	406	TGGATGCTGCTGGGGCGCGCTCTCTCACCGCGCTGGCTGGCTGGCTGGCTGGCTGG 465
Qy	141	AlaProValTyrIleSerGluIlealaTyrProAlaValArgGlyLeuLeuGlySerCys 160
Db	466	GCCCGCGCTACATCTCCGAAATCGCTACCCAGCAGTCCGGGGTGTCTCGGCTCTGT 525
Qy	161	ValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
Db	526	GTGAGCTAACTGCTGCTGGCGAFCCTCTCGCGCGCTTCTGGCGCGCGCGCGCGCG 585
Qy	181	TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
Db	586	TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 645
Qy	201	PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
Db	646	TTTCATGCCCGAGACCCCGCGCTTCTGCTGACTCAGCAGCAGCGCGCGCGCGCGCG 705
Qy	221	AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProIleGlyAlaGlu 240
Db	706	GCCCTGGGTTCTGTGGGCTCCGAGCAGGCTGGAAAGACCCCGCGCGCGCGCGCGCG 765
Qy	241	GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIleGly 260
Db	766	CAGAGCTTTCACCTGGCGCTTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 825
Qy	261	ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
Db	826	GTCTCTCTGATGGCTTCCAGCAGCTGCTGGGGGTCAACCGCGCTCATGTTCATCGAG 885
Qy	281	ThrIlePheGluAlaIlyPheIlyAspSerSerLeuAlaSerValValGlyVal 300
Db	886	ACCATCTTTAAAGAGCCCAAGTTCAGGACAGCAGCGCTGGCTGGCTGGCTGGCTGG 945
Qy	301	IleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArgArgLeu 320
Db	946	ATCCAGGTGCTGTTTCACAGCTGCGCGCTCTCATCATGGAAGAGCAGCGCGCGCGCG 1005
Qy	321	LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
Db	1006	CTCTCTGCTTCTGAGGTGGTTCATGTGTTCAGCAGCGCGCTTTCGGCGCGCTACTTC 1065
Qy	341	IysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer 360
Db	1066	AAGCTGACCCAGGCG 1125
Qy	361	AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu 380

Db 1126 GCACAGCCTGTTGATGCCAGCGTGGGGCTGGCCCTGGCTGGCGGAGCAGCATGTGCCTC 1185
 Qy 381 PheIleAlaGlyPheAlaValAlGlyTrpGlyProIleProTropLeuLeuMetSerGluIle 400
 Db 1186 TTCATCGCCGGTTCGCGGTGGGGCCCAATCCCTGGCTCCTCATGTGAGAGATC 1245
 Qy 401 PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
 Db 1246 TTCCTCTGCATGTCNAGGGCGTGGCGACAGCATCTGGCTCCACCACTGGCTCATG 1305
 Qy 421 AlaPheLeuValThyLysGluPheSerLeuMetGluValLeuAurProTyrGlyAla 440
 Db 1306 GCCTTTCGTCACCAAGAGTTCAGCAGCCTCATGAGAGTTCCTCAGGCCCTATGAGCC 1365
 Qy 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
 Db 1366 TTCCTGGCTTGCCTCCGCTTTCGCAFCATTCAGTGTCCCTTTTCATGTTGTGTCCT 1425
 Qy 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 Db 1426 GAACACTAAGAAAGACACTGGAAACAATCACAGCCCACTTTGAGGGGCGA 1476

RESULT 3
 HSA245937
 LOCUS HSA245937 1873 bp mRNA linear PRI 18-FEB-2000
 DEFINITION Homo sapiens mRNA for glucose transporter (GLUTX1 gene).
 ACCESSION AJ245937
 VERSION AJ245937.1 GI:7018305
 KEYWORDS glucose transporter; GLUTX1 gene.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 AUTHORS Ibberson, M., Uldry, M. and Thorens, B.
 TITLE GLUTX1, a novel mammalian glucose transporter expressed in the
 central nervous system and insulin-sensitive tissues
 JOURNAL J. Biol. Chem. 275 (7), 4607-4612 (2000)
 MEDLINE 20138191
 PUBMED 10671487
 REFERENCE
 AUTHORS Ibberson, M.R.
 TITLE Direct Submission
 JOURNAL Submitted (02-SEP-1999) Ibberson M.R., Institute of Pharmacology
 and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne,
 CH-1005, SWITZERLAND

FEATURES
 source
 1. .1873
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 gene
 1. .1873
 /gene="GLUTX1"
 CDS
 4. .1437
 /gene="GLUTX1"
 /codon_start=1
 /product="glucose transporter"
 /protein_id="CAF75702.1"
 /db_xref="GI:7018306"
 /db_xref="GOA:Q9NY64"
 /db_xref="UniProt/Swiss-Prot:Q9NY64"
 /translation="WTPDEPTEPLGLPPGGAPRRRRVFLAAFAALGPLSFGFAL
 GYSSPAIPLSRAAPPALDDAASWFGAVVTLGAAAGVVLGAWLVDVDRAGKLSLL
 CSVPVFGAVVTAQADVWMLLGRLLRGLACGASIVAPVYISIAIPAVRGLGSC
 VQLMVVGLLALAGWLEWLAIVGCVPLSMMLLWCFMFPETPELLTQHRKQEA
 MAAFLPMSGEQDDPTGAPQSFHALLRQPIYKQFPIIGVLSMAFQQLSGVNAV
 FYAETIFEEKFSDSLASVVGVIQVLFVAALIMDRAGRLLLVLSGVVWFST
 AFGAFKLTGGGPGNSHVAIAPVSAOPVDASVGLAWLVGNMCLFLTAGPAGVGI
 PWLLMSEIFPLHVKVATGICVLTNWLMAFLVTKFSSLMELVRLPYGAFWLASAFICF
 SVLFTLFCVPEIKGILEQIYAFPSR"

ORIGIN Alignment Scores:

Pred. No.: 2.96e-166 Length: 1873
 Score: 2448.00 Matches: 475
 Percent Similarity: 99.79% Conservative: 1
 Best Local Similarity: 99.58% Mismatches: 1
 Query Match: 99.63% Indels: 0
 Gaps: 9
 DB: US-09-886-954A-1 (1-477) x HSA245937 (1-1873)
 Qy 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
 Db 4 ATGACGCCCGAGACCAGAGAAACCCAGCGCTTCTGGGGCTCTCTGGCGGAGCGCG 63
 Qy 21 ProArgGlyArgValPheLeuAlaAlaPheAlaAlaLeuAlaLeuGlyProLeuSerPhe 40
 Db 64 CCCCGGGCCCGCGGCTTCTCGCGCTTCCCGCTGCGCTGGCCCTCAGCTTC 123
 Qy 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 Db 124 GGCCTTGGCGCTCGGCTACAGCTCCCGGCCATCCCTAGCCTGCAGCGCGCGCGCCG 183
 Qy 61 AlaProArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
 Db 184 GCCCGCGCTGGACGACGCGCGCTCTGGTTCGGGGCTGCTGACCTTGGGTGCC 243
 Qy 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
 Db 244 CGCGCGGGGGAGTGTGGCGGCTGGCTGGTGGACCGCGCGGGCCGACAGCTGAGCCTC 303
 Qy 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 Db 304 TTGCTGTGCTCCGCTGCGCTTCCGCGCGCTTTCGCGCTCATCACCGCGCGCCAGACGTG 363
 Qy 121 TrpMetLeuLeuGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 Db 364 TGGATGCTGGGGGGCGCGCTCTCACCGCGCTGGCTGGCTGGGTGGTTCCTCCCTAGTG 423
 Qy 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
 Db 424 GCCCGGCTTACACTCCGAATCGCTTACCAGCAGTCCGGGGGTTCCTCGCTCCTGT 483
 Qy 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
 Db 484 GTGACGCTAATGGTGTGCTGGCATCTCCCTGGCTACCTGGCAGGCTGGGTGTGGAG 543
 Qy 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 Db 544 TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 603
 Qy 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
 Db 604 TTCATGCCCGAGACCCCGCGCTTCTGCTGACTCAGCACAGCGCGCCAGAGGCCATGSC 663
 Qy 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu 240
 Db 664 GCCCTGGCTTCTGTTGGGGCTCCGAGCAGGGCTGGGAAGACCCCGCCATCGGGGCTGAG 723
 Qy 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleGly 260
 Db 724 CAGAGCTTTCACCTGGCGCTGCTGGCGAGCGCGCGCAFCATCAAGCCCTTTCATCAICGC 783
 Qy 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
 Db 784 GTCCTCCCTGATGGCTTCCAGCAGCTGTGGGGGTCAACCGCGCTCATGTTCTATGACAG 843
 Qy 281 ThrIlePheGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVal 300
 Db 844 ACCATCTTGAAGAGGCCCAAGTTCAGGACAGCAGCCTGGCTGGCTGGCTGGCTGGCTG 903
 Qy 301 IleGlnValLeuPheThrAlaValAlaAlaLeuLeuMetAspArgAlaGlyArgArgLeu 320
 Db 904 ATCCAGGTGCTGTTCCAGAGCTGTGGCGCTTCTCATGATGGACAGACAGCGCGGAGGCTG 963

QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
 Db 964 CTCCTGGCTTGTGAGGTGGTCAATGGTGTTCAGCACAGAGTGCCTGGCGGCTACTTTC 1023
 QY 341 LysLeuThrGlnGlyProGlyPheGlyAnserSerHisValAlaIleSerAlaProValSer 360
 Db 1024 AAGCTGACCCAGGGTGGCCCTGGCAACTCTCGACAGTGGCCATCTGGGGCTGTCTCT 1083
 QY 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValAlaGlySerMetCysLeu 380
 Db 1084 GCACAGCCTGTGATGCCACGCTGGGCTGGCTGGCTGGCCGTGGCAACATGTGCCCTC 1143
 QY 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIle 400
 Db 1144 TTCATCGCCGCTTGGCGTGGGCTGGGCGCCATCCCTGGCTTCATGTTCAGAGATC 1203
 QY 401 PheProLeuHisValGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
 Db 1204 TTCCTCTGATGTCAAGGGCGTGGCAGAGGATCTGCGTCTCCACCAACTGGCTCATG 1263
 QY 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
 Db 1264 GCCTTCTGTCGACCAAGGATTCAGCAGCCTCATGGAGTCTCAGGCCCTATGGGCC 1323
 QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
 Db 1324 TTCCTGGCTTCCCTGCTTCTGCACTTCTCAGTGTCTTTCATTTTCTGTTCTGTCCT 1383
 QY 461 GluThrLysGlyLysThrLeuGlnIleThrAlaHisPheGluGlyArg 477
 Db 1384 GAAATTAAGGAAGAGACTCTGGAACAATAACACAGCCCATTTTTCAGGGGGCGA 1434

RESULT 4
 AX076667
 LOCUS AX076667 2217 bp DNA linear PAT 06-FEB-2001
 DEFINITION Sequence 1 from Patent WO0104145.
 ACCESSION AX076667
 VERSION AX076667.1 GI:12711198
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Thoren, B., Ibberson, M. and Uldry, M.
 TITLE Glutx polypeptide family and nucleic acids encoding same
 JOURNAL Patent: WO 0104145-A 1 18-JAN-2001;
 University of Lausanne (CH)
 FEATURES
 Location/Qualifiers
 1..2217
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 348..1781
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC28495.1"
 /db_xref="GI:12711198"
 /translation="MTPEDPEETQLLGPPGSAAPRRRFLAAFAAALGPLSFGFAL
 CVPFPAISLQRAAPPAPRDLDAASWFGAVTLGHANAGVLLGGMLVDRAGKLSILL
 GSVFVAGFAVITAQDQVWMLLGGRLTLGACVAVSLVAVPYISEITAYPAVRLGLSC
 VOLMNVVGLLALYAGLWLEWLVAVLGVVPSLMLLMCFMPEPTFRLLTQRHREA
 MAALRLWMSQEWEDPPIGAEQSFHLALLRQPIKYKPTIIGVSLMAFOLGQVNAV
 FYAETIFERAKPKDSSLASVWVGIQVLFVAALIMDRAGRLILVLSGVVMVST
 AFGAYFLQGGPENSASHVAISAPVSAQVDAVGLAWLAVGNMCLFIAGFAVAGPFI
 PMLMSIEIFLHVKGATGICVLTNWLMAFLVTKBFSLSMEVLRPYGAFWLASAFPCIF
 SVLFTLFCVPEIKGKLEQITAFHEGR"

Best Local Similarity: 99.58% Mismatches: 1
 Query Match: 99.63% Indels: 0
 DB: Gaps: 0
 US-09-886-954A-1 (1-477) x AX076667 (1-2217)
 QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
 Db 348 ATGACGCCCGAGACCCAGAGAAACCCAGCGCTTCTGGGGCTCTCTGGCGCAGCGCG 407
 QY 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaAlaLeuGlyProLeuSerPhe 40
 Db 408 CCCCAGCGCGCGCGCTTCTTCGCGCCCTTTCGCGCTTTCGCGCCCTCAGCTTC 467
 QY 41 GlyPheAlaLeuGlyTyrSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 Db 468 GCGTTTGGCGCTGGCTACAGCTCCCGCGCATCCCTAGCTTGCAGCGCGCGCGCC 527
 QY 61 AlaProArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
 Db 528 GCCCGCGCGCTGGACGACGCGCGCTCTCTGGTTCGGGGCTGTCTGACCCCTGGGTGC 587
 QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
 Db 588 GCGCGGGGGAGTGTGGCGCGCTGGCTGGTGGACCGCGCGCGCGCGCAAGCTGAGCCTC 647
 QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 Db 648 TTGCTGTCTCGTGGCCCTTGTGGCGCGCTTTCGCGCTCATCACCGCGCGCGCGCGCG 707
 QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 Db 708 TGGATGCTCTGGGGCGCGCTCTCTCACCGCGCTGGCTGGGGTGTTCCTCCCTAGTG 767
 QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
 Db 768 GCCCGCGTCTACATCTCCGAAATCGCTACCCAGCAGTCCGGGGTGTCTCGCTCTCTGT 827
 QY 161 ValGlnLeuMetValValValGlyLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
 Db 828 GTGACGTAATGGTGTCTGGCATCTCTTGGCGCTTCTGGCGCGCTGGGTGTGGAG 887
 QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 Db 888 TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 947
 QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
 Db 948 TTCATGCCCGAGACCCCGCGCTTCTTCTGCTGACTCAGCACAGCGCGCGCGCGCGCG 1067
 QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProPheIleGlyAlaGlu 240
 Db 1008 GCGCTGGCGTCTGTGGGGCTTCCGAGCAGGGCTGGGAAGACCCCGCGCGCGCGCGCG 1067
 QY 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleIleIleIleIleIleIleGly 260
 Db 1068 CAGAGCTTTCACCTGGCCCTGTGGCGCAGCCCGCATCTACAAGCCCTTTCATCATCGCG 1127
 QY 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
 Db 1128 GTCTCCCTGATGGCTTCCAGCAGCTGTCTGGGGTCAACCGCGCTCATGTTCATGACAG 1187
 QY 281 ThrIlePheGluAlaLysPheLysAspSerSerLeuAlaSerValValValGlyVal 300
 Db 1188 ACCATCTTTCAGAGGCCAAGTTCAGGACAGCGCTGGCTGGCTGGCTGGCTGGCTGGCT 1247
 QY 301 IleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArgLeu 320
 Db 1248 ATCCAGGTGCTGTTCACAGCTGTGGCGCTCTCATCATGACAGCAGCGCGCGCGCGCG 1307
 QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
 Db 1308 CTCCTGGTCTTGTCAAGGTGTGCTCATGTGTTTCAGCACAGTGCCTTTCGCGCGCTACTT 1367

ORIGIN
 Alignment Scores: 3.61e-166 Length: 2217
 Pred. No.: 2448.00 Matches: 475
 Score: 99.79% Conservative: 1
 Percent Similarity:

MAALRWLWSEGWEDPPICAEQSFHLALRQPIYKYPFIIGVSLMAFOQLSGVAVM
 FVAETFEAKFKDSSLVVGVIGVLFVAALIMDRAGRLLELVLSGVMVNFST
 AFAGYFKLTCFPGNSHVAISAPVSAQVDAVSGFLAVLAVGSMCLFIAGFAVWGPI
 PWLLMSEIFPLHVKGATGICVLTNWMMAFLVTKYFSSLEMLVLRPIYGAFWLASAFCLF
 SVLFTFSCVPEFTKGLQITAFEGR"

ORIGIN

Alignment Scores: 3.75e-166 Length: 1508
 Pred. No.: 2445.00 Matches: 475
 Score: 99.58% Conservative: 0
 Percent Similarity: 99.58% Mismatches: 2
 Best Local Similarity: 99.58% Indels: 0
 Query Match: 99.51% Gaps: 0
 DB: 9

US-09-886-954A-1 (1-477) x HSA17801 (1-1508)

QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
 Db 27 ATGACGCCCGAGGACCAGAGAAACCAGCCGCTTCTGGGGCTCTGGCGGCGGCG 86
 QY 21 ProArgGlyArgValPheLeuAlaAlaPheAlaAlaAlaLeuGlyProLeuSerPhe 40
 Db 87 CCGCGCGCGCGCGCGCTTCTCCCGCGCTTCCCGCGCTTCCCGCGCGCGCGCGCTTCC 146
 QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 Db 147 GCCTTCGGCTCGGCTACAGCTCCCGCGCATCCCTAGCCTGCAGCGCGCGCGCGCG 206
 QY 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
 Db 207 GCGCGCGCGCGCGCGCGCTTCTCCCGCGCTTCCCGCGCTTCCCGCGCGCGCGCGCTTCC 266
 QY 81 AlaAlaGlyGlyValLeuGlyTyrLeuValAspArgAlaGlyArgGlyLeuSerLeu 100
 Db 267 GCGCGCGCGCGCGCGCGCTTCTCCCGCGCTTCCCGCGCTTCCCGCGCGCGCGCGCTTCC 326
 QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 Db 327 TTGCTGCTCCGCTGCGCTTCCCGCGCTTCCCGCGCTTCCCGCGCGCGCGCGCTTCC 386
 QY 121 TrpMetLeuLeuGlyArgLeuLeuThrGlyLeuAlaCysGlyGlyValAlaSerLeuVal 140
 Db 387 TGGATGCTGCTGGGGCGCGCTTCCCGCGCTTCCCGCGCTTCCCGCGCGCGCGCGCTTCC 446
 QY 141 AlaProValTyrIleSerGluIleAlaValAlaValArgGlyLeuLeuGlySerCys 160
 Db 447 GCGCGCGCGCGCGCGCGCTTCCCGCGCTTCCCGCGCTTCCCGCGCGCGCGCGCTTCC 506
 QY 161 ValGlnLeuMetValValGlyIleLeuLeuAlaIleLeuAlaGlyTrpValLeuGlu 180
 Db 507 GTGCAGCTAAATGGTCTCGTGGCATCTCTGGCGCTTCCCGCGCTTCCCGCGCGCGCGCG 566
 QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 Db 567 TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 626
 QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGlnAlaMetAla 220
 Db 627 TTCATGCCCGAGACCCCGCGCTTCCCGCGCTTCCCGCGCTTCCCGCGCGCGCGCGCGCG 686
 QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProIleGlyAlaGlu 240
 Db 687 GCGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 746
 QY 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIleGly 260
 Db 747 CAGAGCTTTCACCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 806
 QY 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
 Db 807 GTCTCCCTGATGGCTTCCAGCAGCTGTCGGGGGTCAACCGCGCTCATGTTCATGCAGAG 866

341 LysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer 360
 1368 AAGCTGACCCAGGGTGGCCCTGGCAACTCTCCGACGTCGGCCATCTCGCGCGCTCTCT 1427
 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu 380
 1428 GCACAGCCTGTTGATGCCAGCGTGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1487
 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIle 400
 1488 TTCATGCCCGCTTTCGGCTGGCTGGCGCCCACTCCCTGCTCTCATGTCAGAGATC 1547
 401 PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
 1548 TTCCTCTGATGTCAGCGCTGGCGCACAGGCACTGGCTCTCCACCAACTGGCTCATG 1607
 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
 1608 GCCTTCTCGTGACCAAGAGTTCAGCAGCCTCATGGAGGCTCTCAGGCCCTCATGGAGCC 1667
 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
 1668 TTCCTGCTGGCTCCGCTTTCGATCTTCAGTCTTTCAGTCTTTCAGTCTTTCAGTCTTTC 1727
 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 1728 GAAATTAAGGAAGACTCTGGAACAAMATCACAGCCCAATTTTGGGGGGCGA 1778

RESULT 5
 HSA17801
 LOCUS Homo sapiens mRNA for glucose transporter 8 (GLUT8 gene). PRI 13-MAY-2000
 DEFINITION
 ACCESSION Y17801
 VERSION Y17801.1 GI:7688145
 KEYWORDS glucose transporter 8; GLUT8 gene.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
 Doege, H., Schurmann, A., Bahrenberg, G., Brauers, A. and Joost, H.G.
 GLUT8, a novel member of the sugar transport facilitator family
 with glucose transport activity
 J. Biol. Chem. 275 (21), 16275-16280 (2000)

JOURNAL 20283667
 MEDLINE 10821868
 PUBMED
 REFERENCE 2 (bases 1 to 1508)
 Joost, H.G.
 Direct Submission
 Submitted (16-JUL-1998) H. Joost, Institute of Pharmacology and
 Toxicology, Technical University Aachen, Wendlingweg 2, D-52057
 Aachen, FRG

FEATURES
 source 1..1508 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="testis"
 1..1508 /gene="GLUT8"
 27..1460 /gene="GLUT8"
 /codon_start=1
 /product="glucose transporter 8"
 /protein_id="CAB89809.1"
 /db_xref="GI:7688146"
 /db_xref="GOA:Q9NY64"
 /db_xref="UniProt/Swiss-Prot:Q9NY64"
 /translation="MTPEDEETQPLILPGGSGAPRRGRVFLAAFAALGPLGFGFAL
 GYSPALPSLQRAAPPAPRLDDAASWFGAVVTLGRAAGVGLGMLVDRAGKLSLLL
 CVPFVAGFAVITAAQVDMLLGRLITGLACGVASLVAIVYISEIAYPAVRGLLGGC
 VQLMVMVGIILAYLAGVLEWRMLAVLGVCFPPSLLMLLMLMCFMPEFTFRFLLTQRRQEA

QY 281 ThrIlePheGluCiuAlaLysPheLysApsSerSerLeuAlaSerValValGlyVal 300
 Db ACCATCTTTGAAGAGCCCAAGTTCAAGGACAGCAGCCTCGCTCGTGGGTGC 926
 QY 301 IleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArgIleu 320
 Db AYCAGGCTGTTCACAGCTGGCGGCTCTCATCATGGACAGACAGCGGGAGGCTG 986
 QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaIlePhe 340
 Db CTCTCGTCTTGTCAAGTGTGGTCAATGGTTCAGCACAGATGGCTTCGGCGCTACTTC 1046
 QY 341 LysLeuThrGlnGlyGlyProGlyAenSerSerHisValAlaIleSerAlaProValSer 360
 Db AAGCTACCCAGGCTGGCCCTGGCAACTCTCGCACCTGGCCACTCTCGGGCTGTCTCT 1106
 QY 361 AlaGlnProValAspIleSerValGlyLeuAlaIleThrLeuAlaValGlySerMetCysLeu 380
 Db GGACAGCCTGTGATGCCAGCGTGGGCTGGCCCTGGCTGGCGTGGCGACAGATGGCTTC 1166
 QY 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProIleProIleLeuMetSerGluIle 400
 Db TTYCATGCCGGCTTTCGGTGGCTGGGCGCCATCCCTGGCTCTCATGTCCAGATC 1226
 QY 401 PheProLeuHisValLysGlyValAlaIleThrGlyIleCysValLeuThrAenTrpLeuMet 420
 Db TTYCCCTCTCATGCTCAAGGGCTGGCGACAGGCTCTGGCTCTCAACCACTGGCTCATG 1286
 QY 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTrpGlyAla 440
 Db GGCCTTCTCGTACCAAGAGTTCACAGCCTCATGGAGTCTCTCAGGCTCATGGAGCC 1346
 QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
 Db TTYTGGCTTCCCTCGCTTTCGATCTTCAGTGTCTTTTACATTTTTTCTGTGTCCCT 1406
 QY 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 Db GAAACTAAAGAAAGACTCTGGAACAAATCACAGCCCAATTTTGGGGGGGA 1457

RESULT 6
 AY208940
 LOCUS 2073 bp mRNA linear MAM 01-JAN-2004
 DEFINITION Bos taurus glucose transporter 8 (GLUT8) mRNA, complete cds.
 ACCESSION AY208940
 VERSION AY208940.1 GI:37784545
 KEYWORDS Bos taurus (cow)
 SOURCE Bos taurus
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovinae; Bos.
 REFERENCE 1. (bases 1 to 2073)
 AUTHORS Zhao, F.-Q., Dong, B. and Zheng, Y.
 TITLE Molecular cloning and expression of GLUT8 in bovine tissues
 JOURNAL Unpublished
 REFERENCE 2. (bases 1 to 2073)
 AUTHORS Zhao, F.-Q.
 TITLE Direct Submission
 JOURNAL Submitted (31-DEC-2002) Department of Animal Science, University of Vermont, 219 Terrill, 570 Main Street, Burlington, VT 05405, USA
 FEATURES
 source
 1. .2073
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /issue_type="kidney"
 1. .2073
 /gene="GLUT8"
 96..1532
 /gene="GLUT8"
 /note="bGLUT8"
 /codon_start=1

/product="glucose transporter 8"
 /protein_id="AAP43920.1"
 /db_xref="GI:37784545"
 /translation="MTPEDQRETOLLRPPGSAAPRRREVFLAFAAALPLSPGFAI
 GYSPAIERBRAAPPAHLDEDAASWFGAIYVTLGNAAGGVVIGSEILLDRAGKLSIVL
 CALPFAVAVITAAAGNLMWLLGGRLITGLACIASLAVAPVYISEIAYPEVRLGHSQ
 VMVVTGILLAYLAGWLEWRKWLAVLGCVPFSPMLLLMFCMPETPRFLPSQHRHQEA
 NAAQFLMGYACQWEPEPLGAQDFVLAQLRRPGVYKFFIIGISLMAFQQLSGVNAV
 MAAMQIFIEBAKFDKSSLASVVVGVIOVLFATAALIMDRAGRLILLTLTSGVAVGVPST
 SFACTYFKLITGGSPNSHVDLPAVLSMEADTNYGLAWLAVGSMCLFTAGFAVWGWP
 IPWLLMSIFPLHVKVATGCVLITWPFMFLVYKFFSLSLMEVLRPYGAFWLASAFCI
 FGVLFLLACVPEPKTKLEQITAEFR"

ORIGIN

Alignment Scores:
 Pred. No.: 1..3e-150 Length: 2073
 Percent: 2230.50 Matches: 425
 Score Similarity: 94.35% Conservative: 26
 Best Local Similarity: 88.91% Mismatches: 26
 Query Match: 90.78% Indels: 1
 DB: 4 Gaps: 1
 US-09-886-954a-1 (1-477) x AY208940 (1-2073)
 QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
 Db 96 ATGACGCTTAGGACCAAGAGGAGACCCAGCGCTCTTACGGCCCGCGCGCAGCGCT 155
 QY 21 ProArgGlyArgArgValPheLeuAlaPheAlaAlaAlaLeuGlyProLeuSerPhe 40
 Db 156 CCCCAGCGCCCGCGGCTCTTCTCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTC 215
 QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaPro 60
 Db 216 GGCCTTGGCCCTCGCTACAGCTCCCGGCCATCCCGAGCTGAGGGCCCGCGCCCG 275
 QY 61 AlaProArgLeuAspAspAlaAlaAlaAlaSerTrpPheGlyValValValThrLeuGlyAla 80
 Db 276 GCCCGCACCTCGACGAGACGACGCTCTCGTGGTTCGGGGCCATCGTAGCCCTGGCGCC 335
 QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
 Db 336 CGCGCCGGGGGCTGCTGGCGGCTGGCTTCGGAACCGCGCGGGCGCAAGCTGAGCCTA 395
 QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 Db 396 GTCCCTGGCCCTTCCCTTTCGCGCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTC 455
 QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaAlaSerLeuVal 140
 Db 456 TGGATGCTGCTCGGAGCGGCTGCTCACCGGCTGGCTGGCGCATTCGCTCGCTCGCTG 515
 QY 141 AlaProValTyrIleSerGluIleAlaIleValProAlaValArgGlyLeuLeuGlySerCys 160
 Db 516 GCCCGGCTTATCTCTGAAATTCCTACCTGAGGTGCGAGGGCTGCTCGGCTCTGT 575
 QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaIleValLeuAlaGlyTrpValLeuGlu 180
 Db 576 GTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 635
 QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 Db 636 TGGCGCTGGCTGGCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 695
 QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
 Db 696 TTCATGCTGAGACCCCTCGCTTCTCGCTTCTCGCTTCTCGCTTCTCGCTTCTCGCT 755
 QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProPheIleGluGlu 240
 Db 756 GCCATGAGTCTCTGTGGGCTATGCCAGGGTGGAGAGCCCTCCCTTCCTGGGCTCAG 815
 QY 241 ---GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIleProPheIle 259

```

/db_xref="GI:12711201"
/translation="MSPEPQRTOPLLRSPGARAPGRVFLATFAAALGPLSFGPAL
GYSPFAIRRTAPALRGTAAASWEGAVVTLGAAGAGVLLGWLDRGRKLSILL
CTVPTFAVITAAADVMMGLRLLTGLACVAVSLVAVPIYSEIAYFAVRGLGSC
VQMLVVTGLLAVVAGVWLEWRWLVLCVPPFLMMLLCCMPYPTFRLLTQHOYGA
MAALRLWSEGEWBERPVGABHQFQFQMLRPLLIIGICLWVFOQLUSGVNAI
MFVANTIPEAKTOSGPNSSHVGLLIPISARPADVHLGLAWLAVRSMCLFLTAGFAVWGP
SAPFTYKLTOSGPNSSHVGLLIPISARPADVHLGLAWLAVRSMCLFLTAGFAVWGP
IPWLLMSEIFPLHIKGVATGVCVLTNNWFAFLVTKFENSIMETLRYPGFAWLTAAFCI
LSVLFLLTTFVFPETKRLQIYAHFEGR"

```

```

ORIGIN
Alignment Scores:
Pred. No.: 9,86e-146 Length: 2087
Score: 2162,50 Matches: 408
Percent Similarity: 92,26% Conservative: 33
Best Local Similarity: 85,36% Mismatches: 36
Query Match: 88,01% Indels: 1
Gaps: 1

```

```

US-09-886-954A-1 (1-477) x AX076669 (1-2087)
QY 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
DB 30 ATGTGCGCCGAGGACCCCAAGAAACACAGCCGCTGTTGGGTCGCCGGGGCCAGGGCT 89
QY 21 ProArgGlyArgValPheLeuAlaAlaPheAlaAlaLeuAlaLeuGlyProLeuSerPhe 40
DB 90 CCCGGCGCCCGGGGTCCTTCCGTCACCTTGGCCGCGCCGCTGGGGCCGCTCAGCTTC 149
QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
DB 150 GGCCTTGGCTTGGCTACAGTCCCGCCGATCCAGCCCTGCGGGCCGACCCGCCCT 209
QY 61 AlaProArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
DB 210 GCCCTACGCTCGGAGACACTCCCGCCCTCTGGTTCGGGGCCGCTGAGCCCTGGGG 269
QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgGlyLeuSerLeu 100
DB 270 GCGCAGGGGGCGTGGTGGGGCTGGCTCTGGACCGTGGAGGGCCGCAAGCTGAGGCTC 329
QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
DB 330 CTGCTCTGACCCGCTTCCGTCACCGGTTTCTGTATCACCGGGCCCGGGGATGTG 389
QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
DB 390 TGGATGCTCGGGAGCCGCTCTCCACCGCCCTAGCCTGGGAGTCCGCTCATTAGTG 449
QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
DB 450 GCACCGGCTATATCTCGAAATCGCTTACCAGCCGCTTCGAGGACTGCTCGGCTCTGT 509
QY 161 ValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
DB 510 GTGAGCTGATGTTGTCACTGGCACTCTTGGCTACGTCGGCAGGCTGGGCTCTAGAA 569
QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
DB 570 TGGCGCTGTTGGCTGCTGGGCTGTGTGCCCGCCACCCCTCATGCTGCTCATGTGC 629
QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgGlnGlnAlaMetAla 220
DB 630 TACATGCCCGAGACCCACCGCTTCTCTGTAGCAACACACAGTACCAGAGGCTATGGCT 689
QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu 240
DB 690 GCCCTGGCTTCTGTGGGCTCTGAGGAGGCTGGGAGAGCCCTGTTGGGCTGAG 749
QY 241 ---GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIle 259
DB 750 CACCAGGGCTTCCAGCTGGCTATGTGAGGGGGCCCTGGTGTGTCCACAGCCCTCATATC 809

```

```

DB 816 CACCAGGACTTCCAGCTGCCAGCCAGCTGCGGGCTCTGCTGCTCAAGCCCTTTCATCATC 875
QY 260 GlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAla 279
DB 876 GGCATCTCCCTGATGCTTCCAGCAGCTGTGACGGCGTCAATGCGCTCATGTTCTACGGC 935
QY 280 GluThrIlePheGluGluAlaIlePheLysAspSerSerLeuAlaSerValValGly 299
DB 936 GAGACCATCTTTCAGGAGGCTAAGTTCAGGACAGCAGCAGCCCTGCTGCTGCTGCGG 995
QY 300 ValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArg 319
DB 996 GTCATCCAGGCTGCTTACTGCCACGGCGCCCTGATCATGACAGAGCTGGGGCAGG 1055
QY 320 LeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyr 339
DB 1056 CTGCTCTGACCTTGTGAGTGTGCTATGCTGTTGAGCAGCCAGCCGCTTGTGTTACCTAC 1115
QY 340 PheLysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProVal 359
DB 1116 TTTCAAGCTGACCGGGCGCCCAAGCAACTCTCCACGTTGGACTCCCGGCGCTGCTGC 1175
QY 360 SerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValIleGlySerMetCys 379
DB 1176 TCCATGGAGCCGCCGATACCAACGTTGGGCTGGCTGGCTGGCGGTGGCAGCATGTGC 1235
QY 380 LeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGlu 399
DB 1236 CTCTTCACTGCGGCTTCTGCTGTTGGCTGGGGCCATCCCGTGGCTCTCATGTTGAG 1295
QY 400 IlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeu 419
DB 1296 ATCTTCCCTCTGCATGTCAGGCGTGGCCACCGGCTGCTGGCTCTCCACCAACTGGTTC 1355
QY 420 MetAlaPheLeuValThrLysGlyPheSerSerLeuMetGluValLeuArgProTyrGly 439
DB 1356 ATGCGCTTCTGTTGACCAAGAGTTCAGAGTCTCATGGAGTGTCTCAGGCGCTTACGG 1415
QY 440 AlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal 459
DB 1416 GCCTTCTGGCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1475
QY 460 PrbGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
DB 1476 CCTGAGACCAAGGAGACTTTGGAGCAATCACAGCCCAATTTGAGGGGGCA 1529

```

```

RESULT 7
AX076669 2087 bp DNA linear PAT 06-FEB-2001
LOCUS Sequence 3 from Patent WO0104145.
DEFINITION AX076669
ACCESSION AX076669
VERSION AX076669.1 GI:12711200
KEYWORDS
SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1
AUTHORS Thorens, B., Ibberson, M. and Uldry, M.
TITLE Glutx polypeptide family and nucleic acids encoding same
JOURNAL Patent: WO 0104145-A 3 18-JAN-2001;
University of Lausanne (CH)
FEATURES
source
location/Qualifiers
1..2087
/organism="Rattus sp."
/mol_type="unassigned DNA"
/db_xref="taxon:101118"
30..1466
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC28496.1"

```

CDS

```

/mol_type="mRNA"
/db_xref="taxon:10116"
1..2087
/gene="GLUTX1"
30..1486
/gene="GLUTX1"
/codon_start=1
/product="glucose transporter"
/protein_id="CA718605"
/db_xref="GI:7018605"
/db_xref="COA:Q9JUZ1"
/db_xref="UniProt/Swiss-Prot:Q9JUZ1"
/translation="MSPEDPQETPLRSPGARAFGRVFLATFAALGLPGLSFGFAL
CVSPFAIPLRRTAPPALRGLDTAASWFGAVTGLAAAGVGLGGMLLDRAGRKLSLL
GTVPTVGTAVTAAARDVMMLGRLTGLAGVAVLAVPVVSETAYPAVRGLGSC
VQMVVVTGLLAVVAGWLEWELAVLGCVPPTMLLMKVMPEPRFLVTOHQOEA
MVALRPLWGSSEWEPVGRSHQFQLAMLRPPGVHVKFLIIGICLLMVFQQLSVGNVA
MFIANTIEFAKPFKDSLSASVTGIIQVLTFAVALIMDRAGKLLMLLUSLWVMSFM
SAFGYFKLQDTPSNSHVGLLVPISAEADVHLGLAWLAVGSMCLFAGFAVGGP
IPWLMSEIFPLHIGKGVATGCVLTNWFMAFLVTKFEFNSIMELRIPYGFALTAFCI
LSVLFLLAVPVPKTRLEQITAFHEGR"

```

```

gene
CDS

```

```

ORIGIN
Alignment Scores:
Pred. No.: 9,866-146 Length: 2087
Score: 2162.50 Matches: 408
Percent Similarity: 92.26% Conservatve: 33
Best Local Similarity: 85.36% Mismatches: 36
Query Match: 88.04% Gaps: 1
DB: 10 Gaps: 1
US-09-886-954A-1 (1-477) x RNO245935 (1-2087)

```

```

QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
Db 30 ATGTCGGCCGAGGACCCCGAGAAACACAGCCGCTGTTCGGTTCGGCCGCGCCAGCGGCT 89
QY 21 ProArgGlyArgValPheLeuAlaPheAlaPheAlaAlaAlaLeuGlyProLeuSerPhe 40
Db 90 CCGCGGCGCGCGCGGCTTCCTCCGACCTTCGCGCGCGCGCGCGCGCGCGCGCGCTCAGCTC 149
QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
Db 150 GCGTTCGGCTGGCTACAGCTCCCGCCATCCCGAGCTTCGGCGCGCGCGCGCGCGCGCGCT 209
QY 61 AlaProArgLeuAspAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 80
Db 210 GCCTACGCTCGGAGACACTCGCGGCTTCCTGGTTCGGGGCGCTCGTGACCCCTGGGCGCT 269
QY 81 AlaAlaGlyGlyValLeuGlyGlyTyrLeuValAspArgAlaGlyArgGlyLeuSerLeu 100
Db 270 GCGGCGAGGGGGCTGCTGGCGGCTGGCTTCCTGGACCGCTGAGGGCGCGCGCGCGCGCTC 329
QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
Db 330 CTGCTCTGCACCGCTTCGTCACCGGCTTCGTCACCGGCTTCGTCATCACCGCGCGCGCGGATGTG 389
QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
Db 390 TGGATGCTCGTTCGGAGCGCGCTTCCTACCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCGCT 449
QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
Db 450 GCACCGGCTATATCTCGGAATCGCTACCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCGCT 509
QY 161 ValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
Db 510 GTGAGCTGATGGTGTCTCCTGGCTTCCTTCGGCTACGTCGCGCGCTTCGCGCGCTTCGCGCGCT 569
QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
Db 570 TGGCGCTGGTGGCTGTGTGGCTGTGTGGCTGTGTGGCTGTGTGGCTGTGTGGCTGTGTGGCTGTGTGGCT 629

```

```

QY 260 GlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAla 279
Db 810 GGCATTGGCTCATGCTTCCAGCAGCTGTCCAGGGTCAACGCCCATGTTCTATGCC 869
QY 280 GluThrIlePheGluGluAlaIlePheLysAspSerSerLeuAlaSerValValGly 299
Db 870 AACACCATCTTTGAGAGGCCAAGTTCAAGGACAGAGCGCTGGCTTCGTCATGTTGGCG 929
QY 300 ValIleGlnValLeuPheThrAlaValAlaAlaLeuMetAspArgAlaGlyArgArg 319
Db 930 ATCACTCAGGTTCTGTTCACTGCTGTGGCGGCTCATATGACAGAGAGCGCGCGGCAAAA 989
QY 320 LeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyr 339
Db 990 CTGCTTCGGCTTGTCCGGTGTGATCATGGTTTTTCAGCATGAGCGCTTCGTCATCTAC 1049
QY 340 PheLysLeuThrGlnGlyProGlyAnSerSerHisValAlaIleSerAlaProVal 359
Db 1050 TTCMAACTGACCCAGAGTGGCCCCAGCAACTCTCCCATGTAGGCCCTCTGGTGGCCCCATC 1109
QY 360 SerAlaGlnProValAspAlaSerValGlyLeuAlaTyrLeuAlaValGlySerMetCys 379
Db 1110 TCCGCAGAGCTGCTGATGTTCACTGGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1169
QY 380 LeuPheIleAlaGlyPheAlaValGlyTyrGlyProIleProIleProIleProIleProIle 399
Db 1170 CTCTTCATCGCTGTTTTGCAGTAGGCTGGGACCCATCCCTGGCTTCATGTCAGAG 1229
QY 400 IlePheProLeuHisValLysGlyValAlaIleThrGlyIleCysValLeuThrAsnTrpLeu 419
Db 1230 ATCTCCCTCTGCACATCAAGGGTGTGGCTACCGGGCTGTGTCTCCCACTGGTTC 1289
QY 420 MetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGly 439
Db 1290 ATGGCTTCTGCTCACCCTGCTTCTGATCTCCTCAGCGCTTCCTTCACCGCTCACCTTTGTC 1409
QY 440 AlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal 459
Db 1350 GCTTCTGCTCACCCTGCTTCTGATCTCCTCAGCGCTTCCTTCACCGCTCACCTTTGTC 1409
QY 460 ProGluThrLysGlyThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db 1410 CCTGAGACTAAAGGACGACTCTGGAAACAAATCACAGCCCAITTCGAGGGAGCGG 1463

```

```

RESULT 8
RNO245935
LOCUS RNO245935 2087 bp mRNA linear ROD 18-FEB-2000
DEFINITION Rattus norvegicus mRNA for glucose transporter (GLUTX1 gene).
ACCESSION AJ245935
VERSION AJ245935.1 GI:7018604
KEYWORDS glucose transporter; GLUTX1 gene.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1
Ibberson, M., Uldry, M. and Thorens, B.
GLUTX1, a novel mammalian glucose transporter expressed in the
central nervous system and insulin-sensitive tissues
J. Biol. Chem. 275 (7), 4607-4612 (2000)
MEDLINE 20138191
PUBMED 10671487
REFERENCE
2 (bases 1 to 2087)
Ibberson, M.R.
Direct Submission
Submitted (02-SEP-1999) Ibberson M.R., Institute of Pharmacology
and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne,
CH-1005, SWITZERLAND
FEATURES
Location/Qualifiers
1..2087
/organism="Rattus norvegicus"

```

QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
 Db 630 TACATGCCGAGACCCAGCCCTCTCTGACGCACACACAGTACACGAGGCGCATGGCT 689
 QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAappProProIleGlyAlaGlu 240
 Db 690 GCCCTGGCTTCTCTGGGGCTCTGAGGAGGGCTGGGAAGAGCCCTGTTGGGGCTGAG 749
 QY 241 ---GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIle 259
 Db 750 CACAGGGCTTCCAGCTGCTATGCTGAGGCGCCCTGGTGTCCACAAAGCCCTCATCATC 809
 QY 260 GlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAla 279
 Db 810 GGCATTTGGCTCATGCTTCCAGCAGCTGTGAGGGTCAAGCCCATCAITGTTCTATGCC 869
 QY 280 GluThrIlePheGluAlaLysPheLysPheLysPheLysPheLysPheLysPheLys 299
 Db 870 AACACCATCTTGGAGGAGCCCAAGTTCAAGGACAGCAGCCTGGCTCATCTGGGC 929
 QY 300 ValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArg 319
 Db 930 ATCATCCAGGTTCTGTTCACTGCTGTGGGGCCCTCATCATGGACAGAGCGGGA 989
 QY 320 LeuLeuLeuValLeuSerGlyValMetValPheSerThrSerAlaPheGlyAlaTyr 339
 Db 990 CTGCTTCTGGCTTGTGCGGGTGTGAATGATGTTTTCAGCATGAGCGCCCTTGGTACCTAC 1049
 QY 340 PheLysLeuThrGlnGlyProGlyAanSerSerHisValAlaIleSerAlaProVal 359
 Db 1050 TTTCAACTGACCCAGAGTGGCCCACTCTCCCATGTAGGCTCTCTGGTGGCCATC 1109
 QY 360 SerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCys 379
 Db 1110 TCCGACAGCTGCTGTGATGTTCACTGGGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1169
 QY 380 LeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGlu 399
 Db 1170 CTTTCACTGCTGGTGTGAGTGGCTGGGACCACTCCCTGGCTCTCATGTGACAG 1229
 QY 400 IlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeu 419
 Db 1230 ATCTTCCCTCTGCATCAATCAAGGGTGTGGCTACCGGCTGTGTCTCTCAACCACTGGTTC 1289
 QY 420 MetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGly 439
 Db 1290 ATGGCCCTTCTGGTGAACCAAGAGTTTAAACAGCATCAAGGATCTCTAGACCCCTACCGC 1349
 QY 440 AlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal 459
 Db 1350 GCCTTCTGGCTCACCGCTGCTTCTGTATCTCTCAGGGTCTTTTTCAGGCTCACCTTTGTC 1409
 QY 460 ProGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 Db 1410 CCTGAGACTAAGGACAGACTCTGGAAACAAATCACAGCCCAITTCGAGGGAGCGG 1463

RESULT 9
 MMU17802
 LOCUS
 DEFINITION 1490 bp mRNA linear ROD 13-MAY-2000
 ACCESSION Y17802
 VERSION Y17802.1 GI:7688219
 KEYWORDS Glucose transporter 8; GLUT8 gene.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 AUTHORS Doege, H., Schurmann, A., Bahrenberg, G., Brauers, A. and Joost, H.G.
 TITLE GLUT8, a novel member of the sugar transport facilitator family
 with glucose transport activity
 JOURNAL J. Biol. Chem. 275 (21), 16275-16280 (2000)
 MEDLINE 20283667

10821868
 2 (bases 1 to 1490)
 Joost,H.G.
 Direct Submission
 Submitted (16-JUL-1998) H. Joost, Institute of Pharmacology and
 Toxicology, Technical University Aachen, Wendlingweg 2, D-52057
 Aachen, FRG
 Location/Qualifiers
 1..1490 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /tissue type="testis"
 1..1490 /gene="GLUT8"
 26..1459 /gene="GLUT8"
 /codon_start=1
 /product="glucose transporter 8"
 /protein_id="CAB89815.1"
 /db_xref="GI:7688220"
 /db_xref="GOA:O9JIF3"
 /translation="MSEEDFQETQPLRPPERTTGRVFLASFAALGPLSFGFAL
 GYSPALPISRRRTAPPALRLGDNAASWFGAVVTLGAAAGGILGMLDRGRKLSLL
 CTVPFTGFAVITAAQDVMVLLGGLLTLGACGVASLVAVPYVSEIAYPAVRGLLGS
 VOLMVTGILLAYVAGVLEWRMLAVGCVPTLMLLLCMVPEPFRLLTHOYQEA
 MAALRFLWSEEGHEEPVGAHQGFQALILRERGIYKPLIIGISLWVFOOLSGVNAI
 MFYANSIPEBAKFKDSSLSVTVGIIQVLFVAVALLMDRAGRRLLLALSGVIVFMS
 SAFETYFKLQSLPSNSHVLVPIAAEVDVQVGLAWLAVGSMCLFIAGPAVWGPI
 PWLMSBIFLHVKGVATGICVLTNWFMAFLVTKFESSVMEMLRPIYGFAPWLTAAPCAL
 SVLFTLAVVPEPTKTRILEQVTAHPEGR"
 Alignment Scores:
 Pred. No.: 1.18e-145 Length: 1490
 Score: 2159.00 Matches: 412
 Percent Similarity: 92.47% Conservative: 30
 Best Local Similarity: 86.19% Mismatches: 34
 Query Match: 87.87% Indels: 2
 DB: 10 Gaps: 2
 US-09-886-954A-1 (1-477) x MMU17802 (1-1490)
 QY 1 MetThrProGluAappProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
 Db 26 ATGTCCTCCGAGGACCCCGAGAGACGCGCTGTTCGGCCACCGGAGCCAGGACT 85
 QY 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
 Db 86 CCGCGCGCGCGCGGCTTCTCTGCTTCTCTGCGCGCGCGCTGCGGACCCCTCAGCTTC 145
 QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnAlaAlaProPro 60
 Db 146 GCGTTCGCGCTCGGCTACAGCTCCCGCCATCCCGAGCTGCGGCGCCGACCCCGC 205
 QY 61 AlaProArgLeuAappAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
 Db 206 GCCCTGCCCTCGGAGACAAATCGCGCCCTCTGTTTGGGGCGCTGTCACCTGGCGCT 265
 QY 81 AlaAlaGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgGlyLeuSerLeu 100
 Db 266 GCTGAGGGGGCATACTGGCGGCTGGCTCTCGGACCGTGGAGGCGGCAAGCTGAGCCTC 325
 QY 101 LeuLeuCysSerValPropheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 Db 326 TTGCTCTGACCGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 385
 QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 Db 386 TGGATGCTGCTCGGAGGCGGCTCTCTCACCGGCTTGGCTGGGAGTGGCTCTCAGCTAGTG 445
 QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160

Db 446 GCACCGGTTTACATCTCGGAAATCGCCTACCCAGCTGTCCGAGGACTGTCTGCTCTCTGT 505
 Qy 161 ValGlnLeuMetValValValGlyIleLeuAlaLeuAlaLeuAlaGlyTrpValLeuGlu 180
 Db 506 GTGCAGCTGATGTGTTTCTACTGTCACCTCTCGCCCTCTGTGGCAGGCTGGGTCCTAGAG 565
 Qy 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 Db 566 TDGCGCTGGCTGCCCTGTCTGCTGCTGTGTGCCCTGTGTGCCCTCATCTGCTGTCTCATGTGC 625
 Qy 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
 Db 626 TACATGCCGAAACCCACCGTCTTCTCCACTCAACACCCAGTACCAGGAGGCCATGGCT 685
 Qy 221 AlaLeuArgPheLeuTrpGlySerGlnGlyTrpGluAspProProlleGlyAlaGlu 240
 Db 686 GCCTTGCGCTTCTGTGGGCTCTGAGGAGGCTGGAGAGCCCTGTGTGGGCTGAG 745
 Qy 241 ---GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIlySerProPheIle 259
 Db 746 CACCAGGGCTCCAGCTGGCCCTGTCTGAGCGCCCTGGCATCTACAAGCCCTCATCATC 805
 Qy 260 GlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAla 279
 Db 806 GCATTTCCCTCATGTCTTCACAGCAGCTGTCCAGGGTCAATGCTATCATGTCTATGCC 865
 Qy 280 GluThrIlePheGluAlaLysPheLysAspSerSerLeuAlaSerValValValGly 299
 Db 866 AACAGCATCTTCAGGAGGCAAGTTCAGAGCAAGCAGCCTGGCCCTGGTCACTGTGGGC 925
 Qy 300 ValIleGlnValLeuPheThrAlaValAlaLeuLeuMetAspArgAlaGlyArgArg 319
 Db 926 ATAATCCAGCTCTGTTCACTGTGTGGCCCTCATCATGGACAGAGAGGCGAAGG 985
 Qy 320 LeuLeuLeuValLeuSerGlyValMetValPheSerThrSerAlaPheGlyAlaTyr 339
 Db 986 CTGCTCTGGCTGTGTGGGCTGTGATCATGGTGTTCAGTATGAGTCCCTTGGTACTAC 1045
 Qy 340 PheLysLeuThrGlnGlyProGlyAanSerSerHisValAlaIleSerAlaProVal 359
 Db 1046 TTCMAACTGACCCAGCCCTCCCAAGCACTCTCCCGTAGGCTG---GTGCCCATC 1102
 Qy 360 SerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCys 379
 Db 1103 GCGGCGGAGCCTGTGGATGTCCAAGTGGACTGGCTGGCTGGCTGTAGCAGCATGTGC 1162
 Qy 380 LeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProIleProIleLeuMetSerGlu 399
 Db 1163 CTCTCATTTGCTGGCTTGTGGGCTGGGCTGGGACCACCTCCCTGGCTCTCATGTCCAG 1222
 Qy 400 IlePheProLeuHisValGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeu 419
 Db 1223 ATCTTCCCTCTGCAAGTCAAGGCTGGTGGTACCGGCATCTGTGTCTCAACCAACTGGTTC 1282
 Qy 420 MetAlaPheLeuValThrIlySerSerSerLeuMetGluValLeuArgProTyrGly 439
 Db 1283 ATGGCCCTTTCTAGTACCAAGAGTTTCAGCAGCGCTCATGGAGTGTCCAGACCTACGGT 1342
 Qy 440 AlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal 459
 Db 1343 GCCTTCTGGCTCACCGCTCTCCGCTCTCAGTGTCTTCACTATTCACACTGACCGTTGTC 1402
 Qy 460 ProGluThrIlySerGlyThrLeuGluIleThrAlaHisPheGluIlyArg 477
 Db 1403 CCTGAGACTAAAGGCGAGGACTCTGGAACAAGTCCACAGCCCATTTTCGAGGACGA 1456
 RESULT 10
 AF232061
 LOCUS AF232061 1795 bp mRNA linear ROD 14-JUL-2004
 DEFINITION Mus musculus glucose transporter GLUT8 mRNA, complete cds.
 ACCESSION AF232061
 VERSION AF232061.2 GI:50295426

Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1795)
 Carayannopoulos, M.O., Chi, M.M., Cui, Y., Pingsterhaus, J.M.,
 McKnight, R.A., Mueckler, M., Devaskar, S.U. and Moley, K.H.
 GLUT8 is a glucose transporter responsible for insulin-stimulated
 glucose uptake in the blastocyst
 Proc. Natl. Acad. Sci. U.S.A. 97 (13), 7313-7318 (2000)
 20319023
 10860996
 REFERENCE 2 (bases 1 to 1795)
 Moley, K.H., Carayannopoulos, M.O. and Cui, Y.
 Direct Submission
 Submitted (07-FEB-2000) OB/Gyn, Washington University, 4566 Scott
 Ave, St. Louis, MO 63110, USA
 On Jul 14, 2004 this sequence version replaced gi:8671757.
 COMMENT
 Location/Qualifiers
 1..1795
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strains="129"
 /db_xref="taxon:10090"
 1..1434
 /codon_start=1
 /product="glucose transporter GLUT8"
 /protein_id="AAF78366.1"
 /db_xref="GI:8671758"
 /translation="NSPEDPQTPQLRPPPEARTPRGRVFLASFAAALGPLSPFGFAL
 GYSSPAISRRTAPPALRNDNAASFAGVTVLGAAGGILGGWLLDSGRKLSLL
 CTVPVGTAVITTAARDVWMLGRLTGLACGVALVAPVYISETAYPAVRLGSG
 VQLVYVTTGILLAYVAGVLEWRWLVLCVPTLMLLMLCYMPEFRFLITQYQEA
 MAALRWSEGEWEEPPVAGHQFGQLALRRPGYIKPLIIGISLWVFOQLSGVNAI
 MPYANSIFEAKFDSSLSAVTGIQVLFVAVAVLIMDRAGRMLLALSLGVMVFSM
 SAFGYFKLQTLSPNSHGVLVPAEPDQVQGLAVLAVGSMCLFIAGFAVGMGPI
 PMLMSEIPLHVGATGICVLTWFMFALVTKFSSVWMLRVPYGAFLTAAPFAL
 SVLETLTVVPETKGRLEQVTAHFEGR"
 1435..1795
 3'UTR
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,41e-145 Length: 1795
 Score: 2156.00 Matches: 411
 Percent Similarity: 92.47% Conservative: 31
 Best Local Similarity: 85.98% Mismatches: 34
 Query Match: 87.75% Indels: 2
 DB: 10 Gaps: 2
 US-09-886-954A-1 (1-477) x AF232061 (1-1795)
 Qy 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProGlyGlySerAla 20
 Db 1 ATGAGTCCCGAGGACCCCGAGACTCAGCGCTATTTCGGCCACCGGAAGCCAGGACT 60
 Qy 21 ProArgGlyArgArgValPheLeuAlaPheAlaAlaLeuLeuGlyProLeuSerPhe 40
 Db 61 CCCCAGGCGCCGGCTTCTCTCCCTTCCTCCGCGCCGCTCGGGACCCCTCAGCTTC 120
 Qy 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 Db 121 GCCTTCGGCTGGCTACAGCTCCCGCCATCCCGCCGCTCCCGCCGACCCGACCCCG 180
 Qy 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
 Db 181 GCCTTCGGCTCCGAGACAATTCGGCTCTCTGGTTCGGGGCTCGTGAACCCCTGGCGCT 240
 Qy 81 AlaAlaGlyValLeuGlyValTrpLeuValAspArgAlaGlyArgIlyLeuSerIleu 100
 Db 241 GCTCAGGGGGCATACTGCGCGCTGGCTCTCTGGCCGCTTCCAGCCGCTCAGGCGCAAGCTGAGCCTC 300
 Qy 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120

KEYWORDS SOURCE ORGANISM

Mus musculus (house mouse) Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1795) Carayannopoulos, M.O., Chi, M.M., Cui, Y., Pingsterhaus, J.M., McKnight, R.A., Mueckler, M., Devaskar, S.U. and Moley, K.H.

TITLE GLUT8 is a glucose transporter responsible for insulin-stimulated glucose uptake in the blastocyst

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (13), 7313-7318 (2000)

MEDLINE 20319023 PUBMED 10860996

REFERENCE 2 (bases 1 to 1795) Moley, K.H., Carayannopoulos, M.O. and Cui, Y.

TITLE Direct Submission

JOURNAL Submitted (07-FEB-2000) OB/Gyn, Washington University, 4566 Scott Ave, St. Louis, MO 63110, USA

COMMENT On Jul 14, 2004 this sequence version replaced gi:8671757.

FEATURES Location/Qualifiers 1..1795 /organism="Mus musculus" /mol_type="mRNA" /strains="129" /db_xref="taxon:10090"

CDS 1..1434 /codon_start=1 /product="glucose transporter GLUT8" /protein_id="AAF78366.1" /db_xref="GI:8671758"

3'UTR ORIGIN Alignment Scores: Pred. No.: 2,41e-145 Length: 1795 Score: 2156.00 Matches: 411

Percent Similarity: 92.47% Conservative: 31 Best Local Similarity: 85.98% Mismatches: 34

Query Match: 87.75% Indels: 2 DB: 10 Gaps: 2

US-09-886-954A-1 (1-477) x AF232061 (1-1795)

Qy 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProGlyGlySerAla 20

Db 1 ATGAGTCCCGAGGACCCCGAGACTCAGCGCTATTTCGGCCACCGGAAGCCAGGACT 60

Qy 21 ProArgGlyArgArgValPheLeuAlaPheAlaAlaLeuLeuGlyProLeuSerPhe 40

Db 61 CCCCAGGCGCCGGCTTCTCTCCCTTCCTCCGCGCCGCTCGGGACCCCTCAGCTTC 120

Qy 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60

Db 121 GCCTTCGGCTGGCTACAGCTCCCGCCATCCCGCCGCTCCCGCCGACCCGACCCCG 180

Qy 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80

Db 181 GCCTTCGGCTCCGAGACAATTCGGCTCTCTGGTTCGGGGCTCGTGAACCCCTGGCGCT 240

Qy 81 AlaAlaGlyValLeuGlyValTrpLeuValAspArgAlaGlyArgIlyLeuSerIleu 100

Db 241 GCTCAGGGGGCATACTGCGCGCTGGCTCTCTGGCCGCTTCCAGCCGCTCAGGCGCAAGCTGAGCCTC 300

Qy 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120

Db 1378 CTTGAGACTAAAGCAGGACTCTGGAAACAAGTCAACAGCCCAATTCAGGGGACCA 1431

RESULT 11
 AX076671 2072 bp DNA linear PAT 06-FEB-2001
 Sequence 5 from Patent WO0104145.
 AX076671
 AX076671.1 GI:12711202

Mus sp.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Thorens,B., Ibberson,M. and Uldry,M.
 Glutx polypeptide family and nucleic acids encoding same
 Patent: WO 0104145-A 5 18-JAN-2001;
 University of Lausanne (CH)
 Location/Qualifiers

FEATURES
 source 1..2072
 /organism="Mus sp."
 /mol_type="unassigned DNA"
 /db_xref="taxon:10095"
 21..1454
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC28497.1"
 /db_xref="GI:12711203"
 /translation="MSFDPQETQPILRPEARTPRRRVFLASFAAALGFLNPLGFAL
 GYSPALPSLRRTAPALRLGDNAASWFGAVTLGAMAGGILGGWLLDRSGKLSLL
 CTVPVTFVAVIYAARDVWMLLGGRLLTGLACGVASLVAIVPVYISEIAYPAVRGLLGGC
 VOLMVTGILLYAVAGVWLRWLVLCVPPPTMLLLCMYMPETPRFLUTHQYQEA
 MAALFLMGSEGEWEPVGAERFQALLRRPFYKLIIGISLMVFQOLGQVNAI
 MFYANSIPEAKFKDSSLASVTGIIQVLFVAALIMDRAGRLILALSGVIMVFSM
 SAFYFKLQSLPSSSHVGLVPIAAEPYDVGVLAWLAVGSMCLFIAGFVWAGSI
 PWLLMSIEFFLHVKGVATGICVLINFWMAFLVTKFNSVMEMLRPIYGAFWLTAAFAL
 SVLFTLVVPPETKGRILEQVTAHFEGR"

CDS
 Alignment Scores: Length: 2072
 Pred. No.: 7.69e-145 Matches: 409
 Score: 2150.00 Conservat: 33
 Percent Similarity: 92.47% Mismatch: 34
 Best Local Similarity: 85.56% Indels: 2
 Query Match: 87.51% Gaps: 2
 DB: 6

US-09-886-954A-1 (1-477) x AX076671 (1-2072)

QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
 Db 21 ATGTCTCCGAGGACCCCGAGGAGCAGCGCGCTATTTCGGGCCACCGAAGCCAGGACT 80
 QY 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
 Db 81 CCCCAGGCGCCGGGCTCTCTCCGCTTCCTTCGCGCGGCTCTGGGACCCCTCAACTTC 140
 QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 Db 141 GCCTTCGGCTCGGCTACAGCTCCCGCCATCCCGAGCTCCCGGCGCCGCGCCCGC 200
 QY 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThreLeuGlyAla 80
 Db 201 GCCCTCGCCTCGGAGACAATCGCGCTCTCTGGGCGGCTCTGTCGCGGCGCCGCGC 260
 QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgGlyLeuSerLeu 100
 Db 261 GCTGAGGGGGGCACTACTGGCGGCTGGCTCTCGGACCGTTCAGGGCGCAAGTGTAGCCTC 320
 QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 Db 321 TTGCTCTGACCGTTCCTTCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 380

Db 301 TTGCTCTGACCGTCCCTTCGTTGACTGGCTTTGCTGTCATCACCGGGCCGGGATGTG 360
 QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 Db 361 TGGATGCTCTCGGAGGCGGCTCTCACCGGCTTTCCTCGGAGTTCGCTCACTAGTG 420
 QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
 Db 421 GCACCGGHTTACATCTCGGAATCGCTTACCAGCTGCCAGGACTGCTCGGCTCCTGT 480
 QY 161 ValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
 Db 481 GTGACGTGATGTTGCTCACTGGCATCTCTGGCTATGTGGCAGGCTGGTCTCTAGAG 540
 QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 Db 541 TGGCGTGGCTGGCGGCTGGGCTGTGGCCGCCCCACCTCACTGCTGCTCATGTGTC 600
 QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGlnAlaMetAla 220
 Db 601 TACATGCCGAGACCCCGTTCCTCTCACTCAACACCAGTACAGGAGCCCATGGCT 660
 QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu 240
 Db 661 GCCTTGGCTTCTGTGGGCTCTGAGGAGGCTGGGAAGCCCGCTGTGGGCTGAG 720
 QY 241 ---GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIle 259
 Db 721 CACCAGGCTTCCAGCTGGCCCTGTGAGCGCCCTGGCATCTCAAGCCCTCATCATC 780
 QY 260 GlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAla 279
 Db 781 GGCATTCCTCATGCTTCCAGCAGCTGTCCAGGCTCAATGCTATCATGTTCTATGCC 840
 QY 280 GluThrIlePheGluAlaLysPheLysAspSerSerLeuAlaSerValValGly 299
 Db 841 AACAGCATCTTCGAGGAGGCGCAAGTTCAGAGCAGCGCTGGCTCGGCTACTGTGGC 900
 QY 300 ValIleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArgArg 319
 Db 901 ATAATCCAGTCTCTGTTCACTGTGGGCGCTCATATGAGCAGCAGGCGGGAAG 960
 QY 320 LeuLeuLeuValLeuSerGlyValMetValPheSerThrSerAlaPheGlyAlaTyr 339
 Db 961 CTGCTCCTGGCTTGTGGGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1020
 QY 340 PheLysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProVal 359
 Db 1021 TTCMACTGACCCAGGCTCCCGCACTCTCCAGTAGGCTG--GTGCCCATC 1077
 QY 360 SerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCys 379
 Db 1078 GCGCGGAGCTGTGGATGTCAGTGGGACTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1137
 QY 380 LeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGlu 399
 Db 1138 CTCPTCAPTGTGCTTGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 1197
 QY 400 IlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeu 419
 Db 1198 ATCTTCCCTGTCATGTCAGGGTGTGGTACCGCATCTGTCTCTCAACCACTGGTTC 1257
 QY 420 MetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGly 439
 Db 1258 ATGGCTTTCAGTGACCAAGAGTTCAGCAGCTCAAGAGAGTCTCAGACCCCTACGGT 1317
 QY 440 AlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal 459
 Db 1318 GCCTTCTGGCTCACCGTGCCTTCTGGCTCTCAGTGTCTTATTCACACTGACCGTGTGC 1377
 QY 460 ProGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477

QY 121 TtpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 Db 381 TGGATGCTGCTCGAGGCGCCCTCTCACCGGCTTGGCTGGAGTGCCTCACATAGT 440
 QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
 Db 441 GCACCGGTTTACATCTCGGAAATCGCCTACCCAGCTGTCGAGGACTGCTCGGCTCTGT 500
 QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTyrValLeuGlu 180
 Db 501 GTGCAGCTGATGGTGTCTACTGGCATCTCTGGCCATATGTGGCAGGCTGGTCTTAGAG 560
 QY 181 TtpArgTyrLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 Db 561 TGGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 620
 QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
 Db 621 TACATGCCGAGACCCAGCTTTCTCTCACTCAACACCCAGTACCAGGAGCCATGGCT 680
 QY 221 AlaLeuArgPheLeuTyrGlySerGluGlnGlyTyrGluAspProPheGlyAlaGlu 240
 Db 681 GCCTTGGCTTCTGTGGGCTCTGAGGAGGCTGGGAAGCCCTGTGGGGCTGAG 740
 QY 241 ---GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIlePheIleIle 259
 Db 741 CACCAGGGCTTCAGCTGGCCCTGCTGAGCGCCCTGGCATCTACAAGCCCTCATCATC 800
 QY 260 GlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAla 279
 Db 801 GGCATTTCCTCATGCTTCCAGCAGCTGTCCAGGGTCAATGCTATCATGTTCTATGCC 860
 QY 280 GluThrIlePheGluGluAlaIlePheLysAspSerSerLeuAlaSerValValGly 299
 Db 861 AACAGCTTCTCAGAGGCCAAGTTCAAGGAGGCTGGAGAGCCCTGTGGGGCTGAG 920
 QY 300 ValIleGlnValLeuPheThrAlaValAlaLeuLeuMetAspArgAlaGlyArgArg 319
 Db 921 ATAATCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 980
 QY 320 LeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyr 339
 Db 981 CTGCTCTGGCTTGTGGGCTGTGATCATGGTGTTCAGTATGAGTGGCTTGGTACCTAC 1040
 QY 340 PheLysLeuThrGlnGlyProGlyAenSerSerHisValAlaIleSerAlaProVal 359
 Db 1041 TTCAACTGACCCAGGCTCCCAACTCTCCCACTGAGGCTG---GTGCCCATC 1097
 QY 360 SerAlaGlnProValAspAlaSerValGlyLeuAlaTyrLeuAlaValGlySerMetCys 379
 Db 1098 GCGGCGAGGCTGTGGATGTCAGTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1157
 QY 380 LeuPheIleAlaGlyPheAlaValGlyTyrGlyProIleProIleLeuLeuMetSerGlu 399
 Db 1158 CTCTTCAATGCTGGCTTGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 1217
 QY 400 IlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrIleLeu 419
 Db 1218 ATCTTCCCTGTCATGTCAGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1277
 QY 420 MetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGly 439
 Db 1278 ATGGCTTCTAGTGAACCAAGGTTCAACAGGCTGTGGAGATGCTCAGACCTACGGT 1337
 QY 440 AlaPheTyrLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal 459
 Db 1338 GCCTTCTGGCTCACCGTCTCTGGCTCTCAGTGTCTTATTCACTGACCGTGTTC 1397
 QY 460 ProGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGlyArg 477
 Db 1398 CCTGAGCTAAAGGAGGACTCTGGAAACAAGTCAACAGCCATTTTCGAGGGACGA 1451

RESULT 12

MMU245936 2072 bp mRNA linear ROD 18-FEB-2000
 Mus musculus mRNA for glucose transporter (GLUTX1 gene).
 AJ245936
 AJ245936.1 GI:7018606
 Glucose transporter; GLUTX1 gene.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Ibberson, M., Uldry, M. and Thorens, B.
 GLUTX1, a novel mammalian glucose transporter expressed in the
 central nervous system and insulin-sensitive tissues
 J. Biol. Chem. 275 (7), 4607-4612 (2000)
 20138191
 10671487
 2 (bases 1 to 2072)
 Ibberson, M.R., Institute of Pharmacology
 and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne,
 CH-1005, SWITZERLAND
 Location/Qualifiers
 1..2072
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 1..2072
 /gene="GLUTX1"
 21..1454
 /gene="GLUTX1"
 /codon_start=1
 /product="glucose transporter"
 /protein_id="CAB75719.1"
 /db_xref="GI:7018607"
 /translation="MSRDPDPTQLPAPPEARTPRRRVFLASFAAALGPNLNFGLFAL
 GYSSPAISLRTPAPALRLGDNRASWFGAVVTLGAAGGILGGWLLDRSGRKLILL
 CVLPVTFVAVTAAARDVMLGGRLGLGACVASLVAFPVYISETAYPAVRLGLGSC
 TQVMVTVYVAYVAGVWLRWLVLCVPEPTMLLMLCMYMPETPRFLLTQHQYEA
 MAAFLNGSEEGEPEPVGAEHQFQALRRPGLIPLIIGISLVFQQLSGVNAI
 MFYANSIFEARFKDLSASVTGIIQVLFVAVAILMDRAGRRLLALLSGLVTFMSM
 SARQTYFKLQSLPSNSHVGIVPAAREPVDVQGLAVGSMCLFIAGFVAGGPI
 PHLMSEIFPLHVKVAVGICVLTINWFVAFVTKFNSWEMELRYPGAFWLTAAFCAL
 SVLFLTVVVPVTKGRTLEQVTAHFGR"
 Alignment Scores:
 Pred. No.: 7,69e-145 Length: 2072
 Score: 2150.00 Matches: 409
 Percent Similarity: 92.47% Conservative: 33
 Best Local Similarity: 85.56% Mismatches: 34
 Query Match: 87.51% Indels: 2
 DB: 10 Gaps: 2
 US-09-886-954A-1 (1-477) x MMU245936 (1-2072)
 QY 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProGlyGlySerAla 20
 Db 21 ATGTCTCCGAGGACCCCGAGGACCGCCGCTATTGCGGCCACCGGAAGCCAGGACT 80
 QY 21 ProArgGlyArgArgValPheLeuAlaPheAlaAlaLeuLeuGlyProLeuSerPhe 40
 Db 81 CCCCAGGCGCGGGTCTTCTCGCTTCTCGCGCGGCTCTGGGACCCCTCAACTTC 140
 QY 41 GlyPheAlaLeuGlyTyrSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 Db 141 GCGTTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 200
 QY 61 AlaProArgLeuAspAlaAlaAlaSerTyrPheGlyValValThrLeuGlyAla 80

LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source
 gene
 CDS
 ORIGIN
 Alignment Scores:
 Pred. No.:
 Score:
 Percent Similarity:
 Best Local Similarity:
 Query Match:
 DB:
 US-09-886-954A-1 (1-477) x MMU245936 (1-2072)
 QY
 Db
 QY
 Db
 QY
 Db
 QY

Db 201 GCCTCGGCTCGAGACAATCGGCTCCTGTTCCGGGGCCCTCGTGACCCCTGGGGCGT 260
 Qy 81 AlaAlaGlyClyValLeuGlyClyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
 Db 261 GCTCAGGGGGATACAGTGGGGCTGGCTCGGACCCCTTCAGGGCCGCAAGCTGAGCCTC 320
 Qy 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 Db 321 TTGCTCTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 380
 Qy 121 TrpMetLeuLeuGlyClyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 Db 381 TGGAGTGTCTCGAGGGCCCTCTCCACCGGCTTGCCTGGAGTGCCTCACTAGTG 440
 Qy 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
 Db 441 GCACCGGTTTACATCTCGGAATCGCTACCCAGCTGTCGAGGACTGCTCGGCTCTGT 500
 Qy 161 ValGlnLeuMetValValGlyLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
 Db 501 GTGAGCTGATGGTGTCTACTGGCATCTCTGGCTAATGTGGAGGCTGGGTCTCTAGAG 560
 Qy 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 Db 561 TGGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 620
 Qy 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
 Db 621 TACATGCCGAGACCCAGCTTCTCTCACTCAACACCCAGTACCCAGGAGCCATGGCT 680
 Qy 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProIleGlyAlaGlu 240
 Db 681 GCCTTGGCTTCTGCTGGGCTCTGAGGAGGCTGGAGAGCCCTGGTGGGCTGAG 740
 Qy 241 --GlnSerPheHisLeuAlaLeuAlaLeuArgGlnProGlyIleTyrLysProPheIleIle 259
 Db 741 CACAGGGCTTCAGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 800
 Qy 260 GlyValSerLeuMetAlaPheGlnLeuSerGlyValAsnAlaValMetPheTyrAla 279
 Db 801 GGCATTTCCCTCATGCTCTCCAGCAGCTGTGAGGGTCAATGCTATCATGTTCTATGCC 860
 Qy 280 GluThrIlePheGluAlaLysPheLysAspSerSerLeuAlaSerValValGly 299
 Db 861 ACAGCATTTTCAGAGGACCAAGTTCAGGACAGCAGCCTGGCTGGCTGGCTGGCTGGCT 920
 Qy 300 ValIleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArg 319
 Db 921 ATAAATCCAGTCTGTACTGCTGGGGCTTCATGAGGAGCCTCACTCATGAGCAGAGCAGG 980
 Qy 320 LeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyr 339
 Db 981 CTGCTCTGGCTTGTGCGGTGTGATCATGTTGTTGATGATGAGTGGCTTGTGATACCTAC 1040
 Qy 340 PheLysLeuThrGlnGlyProGlyAenSerSerHisValAlaIleSerAlaProVal 359
 Db 1041 TTCAACTGACCCAGGCTCCCGCACTCTCCAGTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1097
 Qy 360 SerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCys 379
 Db 1098 GCGGGAGCCTGTGGATGTCGAAGTGGGACTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1157
 Qy 380 LeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGlu 399
 Db 1158 CTCTTCACTGCTGGCTTGTGCGGTGGCTGGGACCCATCCCTGGCTCTCTCAATGTCAGAG 1217
 Qy 400 IlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeu 419
 Db 1218 ATCTTCCCTCTGCAATGTCAGGGTGTGGCTACCGGATCTGTCTCAACCACTGGTTC 1277
 Qy 420 MetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGly 439
 Db 1278 ATGGCTTCTAGTACCAAGATTTCAACAGCGTCAATGGAGATGCTCAGACCCCTACCGT 1337

Qy 440 AlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal 459
 Db 1338 GCCCTTGGCTACCCGCTGCCCTTTCGGCTCTAGTGTCTTATTCACACTGACCGTTC 1397
 Qy 460 ProGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 Db 1398 CCTGAGACTAAAGCAGGACTCTGGAACTGACAGCCATTTTCGAGGGACGA 1451

RESULT 13
 AB033418
 LOCUS AB033418.1 2189 bp mRNA linear ROD 11-APR-2000
 DEFINITION Rattus norvegicus glut8 mRNA for glucose transporter 8, complete cds.

ACCESSION AB033418
 VERSION AB033418.1 GI:7592743
 KEYWORDS glucose transporter 8.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 2189)
 AUTHORS Ishibashi,K.
 TITLE Molecular cloning of a new putative glucose transporter
 JOURNAL Published Only in Database (2000)
 REFERENCE 2 (bases 1 to 2189)
 AUTHORS Ishibashi,K.
 TITLE Direct Submission
 JOURNAL Submitted (08-OCT-1999) Kenichi Ishibashi, Jichi Medical School,
 Pharmacology; Minami-kawachi, Kawachi, Tochigi 329-0498, Japan
 (E-mail:kishiba@jichi.ac.jp, Tel:81-285-58-7326,
 Fax:81-285-44-5541)

FEATURES
 source location/Qualifiers
 1..2189
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone_lib="testis"
 1..2189
 /gene="glut8"
 109..1548
 /gene="Glut8"
 /note="Glut8 has 12 transmembrane segments"
 /codon_start=1
 /product="glucose transporter 8"
 /protein_id="BA094383.1"
 /db_xref="GI:7592743"
 /translation="MSPEDPQETQLLRSPGARPGRRVFLATPAAALGPLSGFGL
 GYSPALPSLRRTAPPALRLDGTAAWFGAVVTLGAAQAGVAVLGLDRAGRLSL
 LCTVPEVTGPAVITAAARDVWMLLGGRLTLGACGVASLVAVPVYISEIYPAVRGLLGS
 CVQLMVTGILLAYVAGVLEWRWLVLCVYPTLMLLCCMCPETPEFLAHTOYQE
 AMAALRFJMWSEGEWEPVGAHEGQGLQMLRRPGRVHKPLIIGLICLWVFOQLSGVNA
 IMFYANTIFEAKFKDSLSASVTGIIQVLFVAALIMDRAGKLLALLSGVIMVFS
 NSAGTFLKFTQSGSNVGLVLPISAEPADVHLGLAWLAVGSMCLFIAGFAVWG
 PIPMLMSEIPLHKVATGCVLTFNMFAPLVTKEFNSIMEILRYPYGAFLWTAFC
 ILSVLFTLTFVPEIKRLEQITAFEGR"

gene
 CDS
 Alignment Scores:
 Pred. No.: 1.59e-144 Length: 2189
 Score: 2146.00 Matches: 407
 Percent Similarity: 91.86% Conserved: 33
 Best Local Similarity: 84.97% Mismatches: 37
 Query Match: 87.34% Gaps: 2
 DB: 10

US-09-886-954A-1 (1-477) x AB033418 (1-2189)
 Qy 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
 Db 109 AFTGCGCCGAGGACCCCGAGAAACACAGCCGCTGTTGCGGTCCCGGGCCAGGCT 168

ORIGIN
 Alignment Scores:
 Pred. No.: 1.59e-144 Length: 2189
 Score: 2146.00 Matches: 407
 Percent Similarity: 91.86% Conserved: 33
 Best Local Similarity: 84.97% Mismatches: 37
 Query Match: 87.34% Gaps: 2
 DB: 10

US-09-886-954A-1 (1-477) x AB033418 (1-2189)
 Qy 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
 Db 109 AFTGCGCCGAGGACCCCGAGAAACACAGCCGCTGTTGCGGTCCCGGGCCAGGCT 168

QY 21 ProArgGlyArgValPheLeuAlaPheAlaAlaLeuGlyProLeuSerPhe 40
 Db 169 CCGGGGCGGGGCGGCTTCCTCCAGGCGGCGGCGGCGGCGGCGGCGGCTTCAGCTTC 228
 QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 Db 229 GGCTTCGGGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTC 288
 QY 61 AlaProArgLeuAspAlaAlaSerTyrPheGlyValAlaValThrLeuGlyAla 80
 Db 289 GCCTACGCCTCGGACACTGGGGCTCTGGTTCGGGGCGGCTTCGTGACCTTCGGGGCT 348
 QY 81 AlaAla---GlyGlyValLeuGlyTyrLeuValAspArgAlaGlyArgIleLeuSer 99
 Db 349 CCGGGCGGCGGCGGCGGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTC 408
 QY 100 LeuLeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAsp 119
 Db 409 CTCCCTGCTTCGACCGGCTTCGGTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTC 468
 QY 120 ValTyrMetLeuLeuGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeu 139
 Db 469 GTGTGATGCTTCGGAGCGGCGGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTC 528
 QY 140 ValAlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuGlySer 159
 Db 529 GTGGCACCGGCTTCATATCGGAAATCGGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTC 588
 QY 160 CysValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTyrValLeu 179
 Db 589 TGTTGTCAGCTGATGTTGTCAGTGGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTC 648
 QY 180 GluTyrArgTyrLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMet 199
 Db 649 GAATGGCGCTGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 708
 QY 200 CysPheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMet 219
 Db 709 TGTCATATGCCGAGACCCCGCTTCCTTCAGGCAACCCAGTACCAGGAGGGCCATG 768
 QY 220 AlaAlaLeuArgPheLeuTyrGlySerGluGlnGlyTyrGluAspProIleGlyAla 239
 Db 769 GCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 828
 QY 240 Glu---GlnSerPheHisLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 258
 Db 829 GAGCACCGAGGCTTCAGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 888
 QY 259 IleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyr 278
 Db 889 ATCGGATTTGCTCCTATGCTTCCAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 948
 QY 279 AlaGluThrIlePheGluAlaIlePheIlePheIlePheIlePheIlePheIlePheIlePheIle 298
 Db 949 GCAACACCATCTTCAGGAGGCGCAAGTCAAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1008
 QY 299 GlyValIleGlnValLeuPheThrAlaValAlaAlaLeuLeuMetAspArgAlaGlyArg 318
 Db 1009 GGCATCATCCAGGCTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1068
 QY 319 ArgLeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAla 338
 Db 1069 AAATGCTTTCGGCTTGTGGGCTGATCATGTTTTCAGCATGAGCGCTTCGTTGCTGCTGCTGCTGCTGCT 1128
 QY 339 TyrPheIleLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaPro 358
 Db 1129 TACTTCAAAATGACAGAGTGGCCCGGCAACTCCTCCCATGATGAGGCTTCCTGCTGCTGCTGCTGCTGCT 1188
 QY 359 ValSerAlaGlnProValAspAlaSerValGlyLeuAlaTyrLeuAlaValIleGlySerMet 378
 Db 1189 ATCTCCGAGAGCTGCTGATGTTACCTTCCCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1248
 QY 379 CysLeuPheIleAlaGlyPheAlaValGlyTyrProIleProTyrLeuLeuMetSer 398

Db 1249 TGCCCTTTCATCGCTGGTTTTTGCAGTAGGCTGGGAGCCATCCCTCGCTCCCTCATGTCA 1308
 QY 399 GluIlePheProLeuHisValIleGlyValAlaAlaThrGlyIleCysValLeuThrAsnTyr 418
 Db 1309 GAGATCTTCCTCCTGCACATCAAGGGTGTGGTACCGGGCTGTGTCTCCTCACCACCTGG 1368
 QY 419 LeuMetAlaPheLeuValThrIleGlyPheSerSerLeuMetGluValLeuArgProTyr 438
 Db 1369 TTCATGGCCTTTCGGTACCAGAGTTCACAGATCATGGAGATCCTCAGACCTTAC 1428
 QY 439 GlyAlaPheTyrLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCys 458
 Db 1429 GCGGCTTTCGCTCACCCTGCCTTCTGTATCCTCAGCGCTCCTTTCCAGCTCACCTT 1488
 QY 459 ValProGluThrIleGlyCysLeuGluIleThrAlaHisPheGluGlyArg 477
 Db 1489 GTCCCTGAGACTAAAGGCAGGACTGTGGAAACAATCAGGCCCATTTTCAGGGGACGG 1545
 RESULT 14
 AX191507
 LOCUS AX191507 1461 bp DNA linear PAT 15-AUG-2001
 DEFINITION Sequence 29 from Patent WO0149728.
 ACCESSION AX191507
 VERSION AX191507.1 GI:15209697
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Kato, S. and Kimura, T.
 TITLE Human proteins having hydrophobic domains and dnas encoding these
 JOURNAL Proteins
 Patent: WO 0149728-A 29 12-JUL-2001;
 Protegene Inc. (JP); SAGAMI CHEMICAL RESEARCH CENTER (JP)
 FEATURES
 source Location/Qualifiers
 1..1461
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 61..849
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC51153.1"
 /db_xref="GI:15209698"
 /translation="MPSDEPTEQLPILGPPGGSPRRRFLVLAAPALGLPLSGFPAL
 GYSPAIPLQRPAAPRLDDAAASWFGAVVTLGAAAGVLLGGLVLDVDRGRKLSLL
 CSVPFVAGVITAAQDVMMLGGRLTLGACVAVLACVAVLVPVYISIAIYAVRGLGSC
 VOLMIVGILLAYLAGWVLRWLAVLGCVPPLMLLMCFMPTPRFLLTQRRQEA
 APGLVRCGHGVQHECLRLLQLQADPGWPLQLARGHLGACLCCTAC"
 ORIGIN
 Alignment Scores: 2.49e-102 Length: 1461
 Pred. No.: 1554.50 Matches: 324
 Score: 67.85% Conservative: 1
 Best Similarity: 67.64% Mismatches: 1
 Best Local Similarity: 63.27% Indels: 153
 Query Match: 6
 Gaps: 2
 US-09-886-954A-1 (1-477) x AX191507 (1-1461)
 QY 1 MetThrProGluAspProGluGluThrGlnProLeuLeuLeuGlyProProGlyGlySerAla 20
 Db 61 ATGACGCCGAGGACCCAGAGAAACCCAGCGCTTCGCGGCTCCTGGCGGACGCGG 120
 QY 21 ProArgGlyArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
 Db 121 CCGGCGGCGGCGGCGGCTTCCTCCAGGCTTCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60

181 GCGTTCGGCTCGGTACAGTCCCGGGCCATCCCTAGCTGCGCGCGCGCCGCCCGCC 240
 61 AlaProArgLeuAspAlaAlaSerTrpPheGlyAlaValValValTrpLeuGlyAla 80
 241 GCGCGCGCGCTGAGCGCGCGCGCTCGGTTCGGGGCTGTCGTGACCCCTGGTGC 300
 81 AlaAlaGlyValLeuGlyTrpLeuValAspArgAlaGlyArgLeuSerLeu 100
 301 GCGCGCGGGAGTGTGCGCGCTGCGTGTGTGAGCCGCGCGGGCGAAGCTGAGCCTC 360
 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 361 TTGCTGTCTCCGTTCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCGCTG 420
 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 421 TGGATGCTCTGGGGCGCGCTCTCCACCGCGCTGGCTGGTGGTGGTGGTGGTGGTGG 480
 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
 481 GCGCGCGTCTACATCTCCGAAATCGCTTACCAGCAGTCCGGGGTTCGCGCTTCGT 540
 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
 541 GTGCACTAATGGTCTGCTGCGCATCTCTGGCTACCTGGCTGGCTGGCTGGCTGGCTGG 600
 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 601 TGGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 660
 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgCysGluAlaMetAl 220
 661 TTCATCCCGAGACCCCGCGCTTCCTGCTGACTCAGCAGCGCGCGAGGCGTCTGCT 718
 220 aAlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaG 240
 718 ----- 718
 240 uGlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleG 260
 718 ----- 718
 260 yValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaG 280
 718 ----- 718
 280 uThrIlePheGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVa 300
 718 ----- 718
 300 IleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArgArgLe 320
 718 ----- 718
 320 uLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPh 340
 719 -----CTGGTTCGTGAGTGTGTCATGGTTCAGCAGCGAGTCCCTGGCGCTACTT 774
 340 eLysLeuThrGlnGlyProGlyAsnSerHisValAlaIleSerAlaProValSer 360
 775 CMAGTCCACCGAGTGGCGCTGGCACTCTCCGACAGTGGCCATCTCGCGCGCTGTCTC 834
 360 rAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLe 380
 835 TGCACAGCTTGTGATCCAGCGTGGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 894
 380 upheIleAlaGly-PheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGlu 400
 895 CTTTCATCCCGGAGGTC----- 911
 400 lePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuM 420
 911 ----- 911

420 etAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyA 440
 912 -----CTCAGGCCCTATGGAG 927
 440 laPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValP 460
 928 CCTTCGGCTTGGCTCCGCTTCTCGCATCTTCAGTGTCCCTTTTTCATTTGCTGTGTC 987
 460 roGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 988 CTGAAACTAAAGAAAGACTCTGGAACAATCAAGCCCATTTTGGAGGGCGGA 1040

RESULT 15
 AF321324 1012 bp mRNA linear MAM 02-JUL-2001
 LOCUS Bos taurus glucose transporter 8 (Glut8) mRNA, partial cds.
 DEFINITION
 AF321324
 ACCSSION
 AF321324.1 GI:14582715
 VERSION
 KEYWORDS
 SOURCE Bos taurus (cow)
 ORGANISM
 Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovinae; Bos.
 1 (bases 1 to 1012)
 Augustin,R.; Navarrete-Santos,A. and Fischer,B.
 Direct Submission
 Submitted (13-NOV-2000) Anatomy and Cell Biology, Faculty of
 Medicine, Martin Luther University Halle, Grosse Steinstrasse 52,
 Halle 061097, Germany
 Location/Qualifiers
 1..1012
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 <1..1012
 /gene="Glut8"
 <1..1005
 /gene="Glut8"
 /codon_start=1
 /product="glucose transporter 8"
 /protein_id="AAK69606.1"
 /db_xref="GI:14582716"
 /translation="ISBIATPEVRLGSGVQLMVTGIIILAYLAGVLEWRWLAVLG
 CVAPSMLLMCFMPTPRFLLSQHQEAMAAQFLWYAQGWEPPLGAGHQDFH
 AQRRPGVYKPIIIGISLMAFOQLSGVNVAFYETIPEAKPKDSLSLAVVGVQV
 LFTATALLMDRAGRLLLTLSGVVVFSAFCTYFKLTEGGPSSSHVDLPALVSM
 EAATNVGLAWLAVGNMCLFIFAGFAVGNFIPMLLMSSEIPLHYKVAATGCVLITWF
 MAFLVTKFSSLMELRPRYGAFWLASAPFIFGLVFLTLACVPEYTKGKTLDEQITAFEGR
 "

Alignment Scores: Length: 1012
 Pred. No.: 2,66e-101 Matches: 293
 Score: 1537.50
 Percent Similarity: 92.81% Conservative: 17
 Best Local Similarity: 87.72% Mismatches: 23
 Query Match: 62.58% Indels: 1
 DB: Gaps: 4

US-09-886-954A-1 (1-477) x AF321324 (1-1012)

QY 145 ileSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGlnLeuMet 164
 Db 1 ATCTCGAAATCGCTTACCCTGAGTGGCGAGGCTGCTCGGCTCCTGCTGTCAGCTGATG 60
 QY 165 ValValValGlyIleLeuLeuAlaTyrIleLeuAlaGlyTrpValLeuGluTrpArgTrpLeu 184
 Db 61 GTAGTCACAGGCACTCTCTAGCCCTACCTGGAGGCTGGGCTCGAGTGGCGCTGGCTG 120
 QY 185 AlaValLeuGlyCysValProProSerLeuLeuLeuMetCysPheMetProGlu 204

Db 121 GCCGTGCTGGGCTGTGGGCCCTCTTCAAGCTGCTCATGTGTTTTCATGCGCTGAG 180
 Qy ThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaAlaLeuArgPhe 224
 Db 181 AGCCCTCGCTTCTGCTGCTCAGCAACAGCACCAGAGGCCATGGCCGCAATGACAGTTC 240
 Qy LeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu---GlnSerPhe 243
 Db 241 CTGTGGGGCTATGCCCCAGGGTTGGAGAGGCTCCCTTGGGGCTCAGCACCAGGACTTC 300
 Qy HisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIleGlyValSerLeu 263
 Db 301 CAGCTGGCCAGTGGGGCTCTGTGTCTACAGCCCTTCAAGCCCTTCAATCGGCATCTCCCTG 360
 Qy MetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGluThrIlePhe 283
 Db 361 ATGGCCCTTCCAGCAGTGTTCAGCGTCAATGCCGTCAATGTTCTACGGGAGACCAATCTTT 420
 Qy GluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyValIleGlnVal 303
 Db 421 GAGGAGCTAAGTTCAGGACAGCAGCCTGGCCTCAGTGGTGGGGCTCATCCAGGTG 480
 Qy LeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArgLeuLeuLeuVal 323
 Db 481 CTGTTCACTGCCACCGCGGCCCTGATCATGACAGAGCTGGGCGAAGGCTGCTTTGACC 540
 Qy LeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPheLysLeuThr 343
 Db 541 TTGTCAAGTGTGGTCAATGTTTCAGCACAGGCCCTTTGGTACCTACTTCAAGCTGACC 600
 Qy GlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSerAlaGlnPro 363
 Db 601 GAGGGCGGCCCCAGCAACTCTCCACAGTGGACTCCCGGGCTCGTCTCCATGGAGGCC 660
 Qy ValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeuPheIleAla 383
 Db 661 GCGGATACCAACGTGGGGCTGGCCCTGGCTGGGGTGGGCAACATGTGCTTTCATCGCC 720
 Qy GlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePheProLeu 403
 Db 721 GGCCTCGCTGTGGGCTGGGGCCCAATCCCGTGGCTCTCATGTACAGATCTTCCCTCTG 780
 Qy HisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPheLeu 423
 Db 781 CATGTCAAGGGCTGGCCACCGGCGTCTGGTCTCCCAACCTGGTTCATGGCTTTCG 840
 Qy ValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTrpLeu 443
 Db 841 GTGACCAAGAGTTCAGCAGTCTCATGAGGTGCTCAGGCCCTCAGGGCTTCTGGCTG 900
 Qy AlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluThrLys 463
 Db 901 GCCTTCGCCCTTCGCAATCTGGGTGCTCTTTCACCTGGCCCTGTGCTCCCTGAGACAAA 960
 Qy GlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 Db 961 GGGAGACTCTGGAAACAATCACAGCCCAATTTGAGGGCCGA 1002

Search completed: February 25, 2005, 20:31:23
 Job time : 5130 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 25, 2005, 16:25:41 ; Search time 644 Seconds
(without alignments)
4384.653 Million cell updates/sec

Title: US-09-886-954A-1

Perfect score: 2457

Sequence: 1 MTPEDPETQPLLPGGSA.....CVPETKGTLEQITAHFEGR 477

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-O=/cgn2_1/USPTD_spool_p/US09886954/runat_23022005_154241_13198/app_query_fasta_1.647
-DB=N_Geneseq_1Dec04 -QWTF=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOF=6
-FGAPEXT=7 -YGAPOF=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_16Dec04:*
1: Geneseqm1980s:*
2: Geneseqm1990s:*
3: Geneseqm2000s:*
4: Geneseqm2001as:*
5: Geneseqm2001bs:*
6: Geneseqm2002as:*
7: Geneseqm2002bs:*
8: Geneseqm2003as:*
9: Geneseqm2003bs:*
10: Geneseqm2003cs:*
11: Geneseqm2003ds:*
12: Geneseqm2004as:*
13: Geneseqm2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Table with 5 columns: Result No., Score, Query Match, Length, DB ID, Description. Contains 5 rows of search results.

Table with 13 columns: ID, Score, Query Match, Length, DB ID, Description. Contains 45 rows of search results.

ALIGNMENTS

RESULT 1
ABZ24794
ID ABZ24794 standard; cDNA; 1445 BP.
XX AC ABZ24794;
XX DT 07-APR-2003 (first entry)
XX DE Human solute carrier type 2A nucleic acid.
XX KW Human; solute carrier type 2A; SLC2A; glucose transporter; p53; cancer;
XX KW cytosstatic; gene therapy; gene; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 7..1440
XX FT /*tag= a
XX FT /product= "Human SCL2A"
XX PN WO200298467-A1.
XX PD 12-DEC-2002.
XX PF 03-JUN-2002; 2002WO-US017419.
XX PR 05-JUN-2001; 2001US-0296076P.
XX PR 10-OCT-2001; 2001US-0328605P.
XX PR 15-FEB-2002; 2002US-0357253P.

XX (EXEL-) EXELIXIS INC.

XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;

XX WPI; 2003-2b1283/19.

XX P-PSDB; ABP58364.

XX Identifying candidate p53 pathway modulating agent as therapeutic target for disorders of defective p53 function e.g. cancer, by assaying purified solute carrier family 2 (SLC2A) polypeptide or nucleic acid with a test agent.

XX Disclosure; Page 45; 58pp; English.

XX The present sequence is that of a human solute carrier type 2A (SLC2A) nucleic acid. Genetic screens were designed to identify modifiers of the p53 pathway in Drosophila in which p53 was overexpressed. Human orthologues (polynucleotides and polypeptides) of one such modifier were then identified, including the present nucleic acid. SLC2As are glucose transporter proteins with sugar transporter domains. SLC2A nucleic acids and polypeptides are attractive drug targets for the treatment of pathologies associated with a defective p53 signalling pathway, such as cancer. The invention provides in vitro and in vivo methods of assessing SLC2A function. Modulation of an SLC2A or its binding partners is useful for understanding the association of the p53 pathway and its members in normal and disease conditions and for developing diagnostic and therapeutic modalities of p53-related pathologies. SLC2A-modulating agents that act by inhibiting or enhancing SLC2A expression, directly or indirectly, e.g. by affecting an SLC2A function such as transport or binding activity, can be identified using methods provided. Modulators include small molecules, nucleic acids, antibodies, antisense, oligonucleotides and phosphothioate morpholino oligomers (claimed).

XX SQ Sequence 1445 BP; 182 A; 500 C; 456 G; 307 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,328-204 Length: 1445
Percent Similarity: 2457.00 Matches: 477
Best Local Similarity: 100.00% Conservative: 0
Query Match: 100.00% Mismatches: 0
DB: 8 Indels: 0 Gaps: 0

US-09-886-954A-1 (1-477) x AB224794 (1-1445)

Table with 4 columns: QY, Db, QY, Db. Contains sequence alignment data for various protein domains and regions, including MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProGlySerAla 20, PrbArgGlyArgValPheLeuAlaPheAlaAlaAlaLeuGlyProLeuSerPhe 40, GlyPheAlaLeuGlyTyrSerProAlaLeProSerLeuGlnArgAlaAlaProPro 60, AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80, AlaAlaGlyValLeuGlyTyrLeuValAspArgAlaGlyArgIleuSerLeu 100, LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120, TrpMetLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140, and TGGATGCTGTGGGGGGGGCTTCCCTCACCGGGCTGGCTGGGGTGTGGCTCCCTAGTG 426.

Table with 4 columns: QY, Db, QY, Db. Contains sequence alignment data for various protein domains and regions, including AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160, ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180, TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuLeuMetCys 200, PheMetProGluThrProArgPheLeuThrGlnHisArgArgGlnGlnAlaMetAla 220, AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu 240, GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIleLysProPheIleIleGly 260, ValSerLeuMetAlaPheGlnLeuSerGlyValValAenAlaValMetPheTyrAlaGlu 280, ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVal 300, IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeu 320, LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340, LysLeuThrGlnGlyProGlyAsnSerHisValAlaIleSerAlaProValSer 360, PheIleAlaGlyPheAlaValGlyTyrGlyProIleProTrpLeuLeuMetSerGluIle 400, PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420, AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440, PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460, GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477, and GAAACTAAAGGAAGACTCTGGAAATAATCACAGCCCAATTTTGGAGGGCGA 1437.

RESULT 2

ABZ24792

ID AB224792 standard; cDNA; 1856 BP.

XX

DB 1084 GCACAGCCTTGATGCCAGCGTGCGGCTGGCTGGCTGGCCCTGGCCGAGCATGTGCCTC 1143
 QY 381 PheIleAlaGlyPheAlaValIleTrpGlyProIleProTrpLeuLeuMetSerGluIle 400
 DB 1144 TTCAATCGGGCTTTTGCAGTGCGGGTGGGGGCGCCTCCCTGCTCAGATCCAGAGATC 1203
 QY 401 PheProLeuHisValIleGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
 DB 1204 TTCCTTCATGTCAGGGCTGGCGAGGAGCATCTGGTCCTCCACCACTGGCTCATG 1263
 QY 421 AlaPheLeuValThrLysGluPheSerLeuMetGluValLeuAlaArgProTrpGlyAla 440
 DB 1264 GCCTTTTCCTGACCAGAGAGTTCAGCAGCTCATGAGGAGTTCCTCAGGCCCTATGGAGCC 1323
 QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
 DB 1324 TTTGCGCTTGCCCTCGCTTTCGACATTCAGTTCCTGCTTTTCCACTTTTCTGTGCTCCT 1393
 QY 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 DB 1384 GRAACTAAGGAAAGACTCTGGAACAATCCAGAGCCCATTTTGAGGGGCGA 1434
 RESULT 3
 AAD09552 ;
 ID AAD09552 standard; cDNA; 2080 BP.
 AC AAD09552;
 XX AAD09552;
 DT 10-SEP-2001 (first entry)
 DE Human transporter and ion channel-1 (TRICH-1) cDNA.

Human; transporter and ion channel-1; TRICH-1; vaccine; cystic fibrosis;
 gene therapy; amyotrophic lateral sclerosis; amnesia; muscular dystrophy;
 hypertension; angina; neurological disorder; asthma; bipolar disorder;
 dementia; depression; Alzheimer's disease; epilepsy; mood; arrhythmia;
 Pick's disease; ischaemic cerebrovascular disease; AIDS; anxiety; stroke;
 Huntington's disease; Parkinson's disease; cerebral neoplasm; allergy;
 demyelinating disease; mental disorder; Schizophrenia; polyomyelitis;
 muscle disorder; diabetes mellitus; myocarditis; Grave's disease;
 dermatomyositis; Sjogren's syndrome; immunological disorder; psoriasis;
 rheumatoid arthritis; Sjoen's disease; systemic lupus erythematosus;
 sickle cell anaemia; Wilson's disease; infertility; Cushing's disease;
 scleroderma; pulmonary artery stenosis; nontropic; Addison's disease;
 malabsorption syndrome; hypercholesterolaemia; cancer; ss.
 Homo sapiens.
 XX OS
 XX Key Location/Qualifiers
 FH 46..1479
 FT /*tag= a
 FT /product= "Human TRICH-1 protein"
 XX
 PN W0200146258-A2.
 XX PD 28-JUN-2001.
 XX PF 22-DEC-2000; 2000MO-US035095.
 XX PR 23-DEC-1999; 99US-0172000P.
 PR 14-JAN-2000; 2000US-0176083P.
 PR 21-JAN-2000; 2000US-0177332P.
 PR 28-JAN-2000; 2000US-0178572P.
 PR 02-FEB-2000; 2000US-0179758P.
 PR 10-FEB-2000; 2000US-0181623P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX PA Baughn MR, Burford N, Au-Young J, Lu DAM, Yang J, Reddy R;
 XX PI Lal F, Hillman JL, Azimzai Y, Yue H, Nguyen DB, Yao MG, Gandhi AR;
 XX PI Tang YT, Khan FA;
 XX PD WPI; 2001-418042/44.

DR P-PSDB; AAE04888.
 XX Novel human transporter and ion channel proteins useful for treating and
 PT preventing transport, neurological, muscle and immunological disorders.
 PS Claim 5; Page 140-141; 160pp; English.
 XX The present sequence is transporter and ion channel-1 (TRICH-1) cDNA.
 CC TRICH is used as vaccine. TRICH is useful for treating a disease or
 CC condition associated with decreased expression of functional TRICH, such
 CC as transport disorder including amyotrophic lateral sclerosis, cystic
 CC fibrosis, Becker's muscular dystrophy, Charcot-Marie Tooth disease,
 CC Duchenne muscular dystrophy, angina and hypertension, neurological
 CC disorders including Alzheimer's disease, amnesia, bipolar disorder,
 CC dementia, depression, epilepsy, ischaemic cerebrovascular disease,
 CC stroke, cerebral neoplasms, Pick's disease, Huntington's disease and
 CC Parkinson's disease, demyelinating diseases, mental disorders including
 CC mood, anxiety, Schizophrenia and seasonal affective disorder, muscle
 CC disorder including cardiomyopathy, myocarditis, polyomyelitis,
 CC dermatomyositis, arrhythmias and asthma and immunological disorders
 CC including AIDS, adult respiratory distress syndrome (ARDS), allergies,
 CC anaemia, diabetes mellitus, rheumatoid arthritis, scleroderma, Sjogren's
 CC syndrome, systemic lupus erythematosus and other diseases including
 CC sickle cell anaemia, Wilson's disease, Cushing's disease, Addison's disease,
 CC artery stenosis, Grave's disease, hypercholesterolaemia, cancers
 CC glucose-galactose malabsorption syndrome, hypercholesterolaemia, cancers
 CC psoriasis and viral, bacterial, fungal, helminthic and protozoal
 CC infections. TRICH DNA is useful in gene therapy and in diagnostic
 CC purposes
 XX SQ Sequence 2080 BP; 300 A; 688 C; 632 G; 460 T; 0 U; 0 Other;
 Alignment Scores: Length: 2080
 Pred. No.: 2,14e-204 Matches: 477
 Score: 2457.00 Conservat: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 100.00% Gaps: 4

US-09-886-954A-1 (1-477) x AAD09552 (1-2080)
 QY 1 MetThrProGluAspProGluGluThrClnProLeuLeuGlyProProGlyGlySerAla 20
 DB 46 ATGACGCCCGGAGGACCCAGAGAAACCCAGCGCTTCTGGGGCTTCTGGCGGACGCGC 105
 QY 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaPheLeuAlaValProLeuSerPhe 40
 DB 106 CCCCAGCGCCCGGGCTTCTTCTCGCGCTTTCGCGCTGGCCCTGGGCGCCACTCAGCTTC 165
 QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 DB 166 GGCCTTGGCTCGGCTACAGCTCCCGGCCATCCCTAGCCATGACCGCGGCCGCCCGCC 225
 QY 61 AlaProArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValAlaValThrLeuGlyAla 80
 DB 226 GCCCAGCGCTGGACGACGCGCGCCCTCTTGGTTGGGGCTTCTGCTGACCTGGGTGCC 285
 QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
 DB 286 GCGCGCGGGGAGTCTGCGCGCTTGGTGAGACCCCGCGCGCGCGCGCGCAAGCTAGACCTC 345
 QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 DB 346 TTGCTGTGCTCGTCCCTTCCGCGCCCTTCTGGCGGCTTTCCTGCTGACCTGGGTGCC 405
 QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 DB 406 TGGATGCTCTGGGGCGCCCTCTCCACCGCGCTGGCTGGGTGGTGTTCCTCCCTCCCTAGTGTG 465
 QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuGlySerCys 160
 DB 466 GCCCGGCTTACATCTCCGAATTCGCATCCAGCAGCTACCCAGGAGTTCGCGGGGTTGCTCGGCTCT 525

```

QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaIleLeuLeuAlaIleGlyTrpValLeuGlu 180
Db 526 GTGACGTAATGTCGCTGCTGGCCTACCTCTGGCTTACCTGGCAGGCTGGGTGGTGAG 585
QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
Db 586 TGCCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 645
QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
Db 646 TTCATGCCGAGACCCGGCTTCCTGCTGACTCAGCAGAGGCCGAGGAGGCCATGGCC 705
QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu 240
Db 706 GCCCTGGCTTCCTGTGGGCTCCGAGCAGGCTGGAGACCCTCCCATCGGGCTGAG 765
QY 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTrpIleTrpIleLeuIleGly 260
Db 766 CAGAGCTTTTTCATGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 825
QY 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
Db 826 GTCCTCCCTGATGGCTTCAGCAGCTGTGGGGTCAACGCTCAGTTCATGTCAGAG 885
QY 281 ThrIlePheGluGluAlaIleLysPheLysAspSerSerLeuAlaSerValValGlyVal 300
Db 886 ACCATCTTTGAGAGGCCAAGTTTCAAGGACAGCAGCTGGCTGGCTGGCTGGCTGGCTGG 945
QY 301 IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgLeu 320
Db 946 ATCCAGGTGTGTTTCAGCTGTGGGGCTTTCATCATGGACAGCAGGGGAGGCTG 1005
QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
Db 1006 CTCTGGCTTTTCAGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 1065
QY 341 LysLeuThrGlnGlyProGlyAenSerSerHisValAlaIleSerAlaProValSer 360
Db 1066 AAGTCAACCAGGCTGGCTGGCACTCTCGCAGGTGGCCATCTGGGGCTGTCT 1125
QY 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu 380
Db 1126 GCACAGCTGTTGATGCCAGCTGGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1185
QY 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIle 400
Db 1186 TTCATGCCGCTTTCGCGTGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 1245
QY 401 PheProLeuHisValLysGlyValAlaIleThrGlyIleCysValLeuThrAsnTrpLeuMet 420
Db 1246 TTCCTCTGCATGTCAAGGGCTGGCCAGCAGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1305
QY 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
Db 1306 GCCCTTCTGGTACCAAGAGTTTCAGCAGCTTCATGGAGGCTTCATGGGCTTCATGGG 1365
QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
Db 1366 TCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1425
QY 461 GluThrLysGlyLysThrLeuLeuGlnIleThrAlaHisPheGluIleArg 477
Db 1426 GAAACTAAGGAAAGACTTGGACCAAAATCAGCCCAATTTTGAGGGGGGA 1476

```

```

DE Human solute carrier type 2A nucleic acid 7657680.
XX Human; solute carrier type 2A; SLC2A; glucose transporter; p53; cancer;
KW cytostatic; gene therapy; gene; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FH 4..1437
FT /*tag= a
FT /product= "Human SCL2A"
XX /product= "Human SCL2A"
PN WO200298467-A1.
XX 12-DEC-2002.
XX 03-JUN-2002; 2002WO-US017419.
XX 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.
PR 15-FEB-2002; 2002US-0357253P.
XX (EXEL-) EXELIXIS INC.
XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
PI WPI: 2003-201283/19.
DR P-PSDB; ABP58364.
XX
XX Identifying candidate p53 pathway modulating agent as therapeutic target
XX for disorders of defective p53 function e.g. cancer, by assaying purified
XX solute carrier family 2 (SLC2A) polypeptide or nucleic acid with a test
XX agent.
XX Disclosure; Page 43-44; 58pp; English.
CC The present sequence is that of human solute carrier type 2A (SLC2A)
CC nucleic acids 7657680. Genetic screens were designed to identify
CC modifiers of the p53 pathway in Drosophila in which p53 was
CC over-expressed. Human orthologues (polynucleotides and polypeptides) of
CC one such modifier were then identified, including nucleic acid 7657680.
CC SLC2A are glucose transporter proteins with sugar transporter domains.
CC SLC2A nucleic acids and polypeptides are attractive drug targets for the
CC treatment of pathologies associated with a defective p53 signalling
CC pathway, such as cancer. The invention provides in vitro and in vivo
CC methods of assessing SLC2A function. Modulation of an SLC2A or its
CC binding partners is useful for understanding the association of the p53
CC pathway and its members in normal and disease conditions and for
CC developing diagnostic and therapeutic modalities of p53-related
CC pathologies. SLC2A-modulating agents that act by inhibiting or enhancing
CC SLC2A expression, directly or indirectly, e.g. by affecting an SLC2A
CC function such as transport or binding activity, can be identified using
CC methods provided. Modulators include small molecules, nucleic acids,
CC antibodies, antisense oligonucleotides and phosphothioate morpholino
CC oligomers (claimed)
XX SQ Sequence 1873 BP; 279 A; 621 C; 573 G; 400 T; 0 U; 0 Other;

```

```

Alignment Scores: 1.14e-203 Length: 1873
Pred. No.: 2448.00 Matches: 475
Score:
Percent Similarity: 99.79% Conservative: 1
Best Local Similarity: 99.58% Mismatches: 1
Query Match: 99.63% Indels: 0
DB: Gaps: 0
US-09-886-954A-1 (1-477) x ABZ24793 (1-1873)
QY 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
Db 4 ATGACCGCCGAGACCAGAACCCAGCCGCTTCCTGGGGCTCCCTGGCCGAGCGG 63
QY 21 ProArgGlyArgArgValPheLeuAlaIlePheAlaAlaAlaLeuGlyProLeuSerPhe 40

```

```

RESULT 4
ABZ24793
ID ABZ24793 standard; cDNA; 1873 BP.
XX AC ABZ24793;
XX DT 07-APR-2003 (first entry)

```

1144 TTCAATCGCGGCTTTGGTGGGCGCCATCCCTGGCTCATGTGACAGATC 1203
 401 PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
 1204 TTCCCTCTGCATGTCGAAGGGCGTGGCGACAGGCATCTGGCTCCACCAACTGGCTCATG 1263
 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
 1264 GCCTTTCTGTGACCAAGAGTTCCAGCCTTCAGGAGTCTCAGGCGCCATATGGAGCC 1323
 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
 1324 TTCTGGCTTGCCTCCGCTTTCGATCTTCAGTGTCTTTTTCATTTGTGTGTCCCT 1383
 461 GluThrLysGlyLysThrLeuGluIleThrAlaHisPheGluGlyValArg 477
 1384 GAAATTAAGAAAGAACTCTGGAAACAATATCACAGCCATTTTGGGGGGGA 1434

RESULT 5
 AAF55865
 ID AAF55865 standard; cDNA; 2217 BP.
 XX AAF55865;
 AC
 XX
 DT 17-APR-2001 (first entry)
 XX
 DE Human GLUTX1 coding sequence.
 XX
 KW Human; GLUTX; gene therapy; vaccine; hexose transport modulator;
 KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia; ss;
 KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 XX
 OS Homo sapiens.
 XX
 FN W0200104145-A2.
 XX
 PD 18-JAN-2001.
 XX
 PF 14-JUL-2000; 2000WO-IB001042.
 XX
 PR 14-JUL-1999; 99US-0143907P.
 PR 27-AUG-1999; 99US-0151140P.
 PR 23-FEB-2000; 2000US-0184285P.
 PR 13-JUL-2000; 2000US-00616132.
 XX
 PA (UYLA-) UNIV LAUSANNE.
 XX
 PI Thorens B, Ibberson M, Uldry M;
 XX
 DR WPI; 2001-112615/12.
 DR P-PSDB; AAB66932.
 XX
 PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in the
 PT prevention, diagnosis and treatment of hexose transport disorders, e.g.
 PT ischaemia and diabetes.
 XX
 PS Claim 3; Page 70-71; 124pp; English.
 XX
 CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facultative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is the coding sequence
 CC for human GLUTX1
 XX
 SQ Sequence 2217 BP; 320 A; 750 C; 672 G; 475 T; 0 U; 0 Other;

64 CCCCCGGGCGCGGCTTCTCTCGCGCCCTTCCGCGCTGCGCCCACTCAGCTTC 123
 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 124 GCGTTTGGCGTCCGGTACAGTCCCGCGCCAFCCCTAGCCCTGACAGCGCCGCCCGCC 183
 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
 184 GCCCGCGCTGCGACGCGCGCCCTTCTGTTCCGGGCTGTCGTGACCTGGGTGCC 243
 81 AlaAlaGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
 244 GCGCGGGGGAGTGTGGCGGCTGGCTGCTGAGCCGCGCGCGCGCAAGCTGAGCCTC 303
 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 304 TTGCTGTCTCCGTGCGCTTCTGGCGGCTTGGCGGCTTTCAGTCAACCGCGCGCCAGGAC 363
 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 364 TGGATGCTGCTGGGGGCGCGCTTCTCACCGGCTGGCTGGCTGGCTGGCTGGCTGGCT 423
 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
 424 GCGCGGCTTACATCTCCGAAATPCGCTTACCAGCAGTCCGGGGTGTCTGGCTCCCTGT 483
 161 ValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
 484 GTGCAGCTAATGCTGCTGGCATCTCTGGCTTCTGGCTTCTGGCTTCTGGCTTCTGG 543
 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 544 TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 603
 201 PheMetProGlnThrProArgPheLeuLeuThrGlnHisArgGlnGlnAlaMetAla 220
 604 TTCATGCCCCGAGACCCCGCGCTTCTGCTGACTCAGCAGCGCCAGGAGCCATGGCC 663
 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu 240
 664 GCGCTGGCTTCTGTGGGCTCCGAGCAGGCTGGGAAAGACCCCGCTATCGGGGCTGAG 723
 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleLeGly 260
 724 CAGAGCTTTCACCTGGCTTCTGGCGGCGCCGCGCATCTACAAAGCCCTTCATCGGC 783
 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
 784 GTCTCCCTGATGGCTTCCAGCAGCTGTGGGGTCAACCGCTCATGTTCATATGCGAGAG 843
 281 ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVal 300
 844 ACCATCTTGAAGAGGCGAAATTCAGGACAGCAGCCTGGCTGGCTGGCTGGCTGGCT 903
 301 IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgLeu 320
 904 ATCCAGGTGCTTTCACAGCTGTGGCGGCTTCTCATCATGGACAGAGCAGGCGGAGGCTG 963
 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
 964 CTCCTGGCTTGTCCAGGTGTGTCATGTTGTCAGCAGAGTGTGCTGGCGGCTACTTC 1023
 341 LysLeuThrGlnGlyProGlyValAsnSerSerHisValAlaIleSerAlaProValSer 360
 1024 AAGCTGACCCAGGGTGGCGGCTGGCGAATCTCTGCGACGTGGCCATCTCGGGCGCTCTCT 1083
 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValAlaGlySerMetCysLeu 380
 1084 GCACAGCTGTGTATGCCAGCGTGGGGCTGGCGTGGCTGGCTGGCTGGCGCAACATGTGCTC 1143
 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluLe 400

Alignment Scores:
 Pred. No.: 1,428-203 Length: 2217
 Score: 2448.00 Matches: 475
 Percent Similarity: 99.79% Conservative: 1

Best Local Similarity: 99.58% Mismatches: 1
 Query Match: 99.63% Indels: 0
 DB: 4 Gaps: 0

US-09-886-954a-1 (1-477) x AAF55865 (1-2217)

QY 1 MetThrProGluuAspProGluuThrGlnProLeuLeuGlyProGlyGlySerAla 20
 DB 348 ATGACGCCCGAGGACCCAGAGAAACCAGCCGCTTCTGGGGCTCCTGGCGGCGCGG 407
 QY 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
 DB 408 CCCCAGGCGCGCGCTTCTCGCGCGCTTCTCGCGCGCTTCTCGCGCGCTTCTCGCG 467
 QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 DB 468 GGCCTTGGCGCTGGCTACAGTCCCGCGCCATCCCTAGCTGCGAGCGCGCGCGCGCG 527
 QY 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
 DB 528 GDCGCCGCGCTGACGACGCGCGCGCTCCTCGGCTTGGGGCTGCTGACCCCTGGGTGCC 587
 QY 81 AlaAlaGlyGlyValLeuGlyTyrTrpLeuValAspArgAlaGlyArgGlyLeuSerLeu 100
 DB 588 GGGCGGGGGAGTGTGGCGCGCTGGCTGGTGGACCGCGCGCGCGCGCGCGCGCGCTC 647
 QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 DB 648 TTGCTGTCTCGTGGCTTGGCGCGCTTGGCGCGCTTGGCGCGCTTGGCGCGCGCGCG 707
 QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 DB 708 TGGATGCTGTGGGGGGCGCGCTCCTCACCGCGCTGGCGCTGGCGGTGTGGCTCCTTAGTG 767
 QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuGlySerCys 160
 DB 768 GCCCGGGTCTACATCTCCGAAATCGCTACCCAGCAGTCCGGGGGTGTGGCGCTCTGT 827
 QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
 DB 828 GTACAGTAAATGCTGTGGCGCACTCTCTGGCGCTTGGCGCGCTTGGCGCGCTTGGCG 887
 QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 DB 888 TGGCGTGGCTGGCTGGCTGGCGCGCTGGCGCGCTGGCGCGCTGGCGCGCTGGCGCG 947
 QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
 DB 948 TTCATCCCGAGACCCCGCGCTTCTGTGACTCAGCAGCGCGCGCGCGCGCGCGCGCG 1007
 QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProPheGlyAlaGlu 240
 DB 1008 GCCCTCGGTTCTGTGGGGCTCCGAGAGGGCTGGAGAGACCCCGCGCGCGCGCGCGCG 1067
 QY 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIlePheProPheIleLeuGly 260
 DB 1068 CAGAGCTTTCACCTGGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1127
 QY 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
 DB 1128 GTCTCCCTGATGGCTTCCAGCAGCTGTGGGGGTCAACCGCGCTGATGTATGACGAG 1187
 QY 281 ThrIlePheGluAlaIlePheLeuAspSerSerLeuAlaSerValValGlyVal 300
 DB 1188 ACCATCTTGAAGAGCCCAAGTTCAAGAGCAGCGCGCTGGCTGGTGGTGGGTGTC 1247
 QY 301 IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgLeu 320
 DB 1248 ATCCAGGTGCTGTTCACAGCTGTGGCGCTCTCATCATGGACAGCAGCGCGCGCGCG 1307
 QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
 DB 1308 CTCCTGGTCTTGTGAGGTGTGGTCAATGGTGTTCAGCAGCGAGTGGCTTGGCGCGCTACTTC 1367

QY 341 LysLeuThrGlnGlyProGlyAnSerSerHisValAlaIleSerAlaProValSer 360
 DB 1368 AAGCTGACCCAGGGTGGCGCTGGCAACTCTCGCACGTGGCCATCTCGGGCGCTGTCT 1427
 QY 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu 380
 DB 1428 GCACAGCCTGTTGATGCTGAGCGTGGCGCTGGCGCTGGCGCTGGCGCAACATGTGCTC 1487
 QY 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuMetSerGluIle 400
 DB 1488 TTCATCGCCGGCTTGGCGTGGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1547
 QY 401 PheProLeuHisValIleValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
 DB 1548 TTCCTCTGCAATGTCAGGGCGTGGCGCAGGCACTGGCTCTCACCAACTGGCTCATG 1607
 QY 421 AlaPheLeuValThrIleGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
 DB 1608 GCCTTCTCGTACCAGGAGTTTCAGCAGCTCATGGAGTCTTCAGGCGCTATGGAGCC 1667
 QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
 DB 1668 TTCTGGCTTGGCTTGGCGTGGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1727
 QY 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 DB 1728 GAAATTAAGGAAGACTCTGGAAACAATCACAGCCCAATTTGAGGGGGCGA 1778

RESULT 6
 AAF55866
 ID AAF55866 standard; cdna; 2087 BP.
 XX
 AC AAF55866;
 XX
 DT 17-APR-2001 (first entry)
 XX
 DE Rat GLUTX1 coding sequence.
 XX
 KW Rat; GLUTX; gene therapy; vaccine; hexose transport modulator;
 KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia; ss;
 KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 XX
 OS Rattus sp.
 XX
 PN WO200104145-A2.
 XX
 PD 18-JAN-2001.
 XX
 PF 14-JUL-2000; 2000WO-IB001042.
 XX
 PR 14-JUL-1999; 99US-0143907P.
 PR 27-AUG-1999; 99US-0151440P.
 PR 23-FEB-2000; 2000US-0184285P.
 PR 13-JUL-2000; 2000US-00616132.
 XX
 PA (UYLA-) UNIV LAUSANNE.
 XX
 PI Thorens B, Ibberson M, Uldry M;
 XX
 DR WPI; 2001-112615/12.
 DR P-PSDB; AAB66933.
 XX
 XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in the
 PT prevention, diagnosis and treatment of hexose transport disorders, e.g.
 PT ischaemia and diabetes.
 XX
 PS Claim 3; Page 71-73; 124pp; English.
 XX
 CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
 CC AAB6932-AAB6941). The GLUTX proteins are related to the facultative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and

CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is the coding sequence
 CC for rat GLUTX1
 XX
 SQ Sequence 2087 BP; 355 A; 673 C; 580 G; 479 T; 0 U; 0 Other;

Alignment Scores: Length: 2087
 Pred. No.: 2162.50 Matches: 408
 Score: 92.26% Conservatives: 33
 Percent Similarity: 85.36% Mismatches: 36
 Best Local Similarity: 88.01% Indels: 1
 Query Match: 4 Gaps: 1
 DB:

US-09-886-954A-1. (1-477) x AAF55866 (1-2087)

QY	1	MetThrProGluAspProGluThrGlnProLeuLeuGlyProGlySerAla	20
DB	30	ATGTCCCGGAGACCCCGAGAAACACAGCGCTGTTGCGGTCCCGGGCCAGGGCT	89
QY	21	ProArgGlyArgValPheLeuAlaPheAlaAlaLeuGlyProLeuSerPhe	40
DB	90	CCCGGGCCCGCGGGTCTTCTTGCACCTTTCGGCCCGCCCTCGGGCCGCTCAGCTC	149
QY	41	GlyPheAlaLeuGlyTyrSerProAlaIleProSerLeuGlnArgAlaAlaProPro	60
DB	150	GGCTTCGGCTTGGCTACAGTCCCGCCGCAATCCCGAGCTCGGGCCAGCCAGCCCT	209
QY	61	AlaProArgLeuAspAlaAlaLeuSerTrpPheGlyAlaValValThrLeuGlyAla	80
DB	210	GCTCTGCGGAGACTGCGGCTCTCGTTCGGGGCCGTGCTGACCCCTGGGCGCT	269
QY	81	AlaAlaGlyValLeuGlyTyrLeuValAspArgAlaGlyArgIleLeuSerLeu	100
DB	270	CGGGCCGGGGGCTGCTGGGGGCTGGCTCTCGACCGCTGAGGGGCAAGCTGAGCTC	329
QY	101	LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal	120
DB	330	CTGCTGCGCCGCTTCTCGTACCGGCTTTCGTCATCACCCCGGGCCGGGATGTG	389
QY	121	TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal	140
DB	390	TGGATGTGTCGGAGCGCGCTCTCCCGGCTAGCCCTGCGGAGTGGCTCATTAGTG	449
QY	141	AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys	160
DB	450	GCACCGGCTATATCTCGGAATCGCTACCCAGCCGTTCCGAGGACTGCTCGGCTCTGT	509
QY	161	ValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTyrValLeuGlu	180
DB	510	GTGCAGCTGATGGTGTCTCACTGGGATCTCTTGGCCCTACGTGGGAGGCTGGTCCAGAA	569
QY	181	TrpArgTyrLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys	200
DB	570	TGGCGCTGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT	629
QY	201	PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla	220
DB	630	TACATGCCGAGACCCACCGTCTCTCTGACGCAACACACAGTACCAGAGGCGCATGGCT	689
QY	221	AlaLeuArgPheLeuTyrGlySerGluGlnGlyTyrGluAspProProIleGlyAlaGlu	240
DB	690	GGCTTCGGCTTCTGTGGGCTCTGAGGAGGCTGGGAAGAGCCCGCTGGTGGGCTGAG	749
QY	241	--GlnSerPheIleLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIle	259
DB	750	CACCGGGCTTCCAGCTGGCTATGCTGAGGGCCCTGGTGTCCACAGCCCTCATCATC	809
QY	260	GlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAla	279
DB	810	GGCATTTGGCTCATGTTCTCCAGCAGCTGTCCAGGGGTCAGGGGTCATGTTCTATGCC	869

QY	280	GluThrIlePheGluGluAlaAlaAlaPheLeuAspSerSerLeuAlaSerValValValGly	299
DB	870	AAACACCACTTTGAGGAGCCCAAGFTTCAAGGACAGACCTTGGCTCGGTCACTGGGCG	929
QY	300	ValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArg	319
DB	930	ATCATCCAGGTTCTGTTCACTGCTGTGGGGCCCTCATCATGGACAGACGGCGGAAA	989
QY	320	LeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyr	339
DB	990	CTGCTTTCGGCCCTTGTCCGGGTGTGATCATGTTTTTCAGCATGAGGCGCTTTGGTACCTAC	1049
QY	340	PheLysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProVal	359
DB	1050	TTCAAACCTGACCAGAGTGGCCCGACCACTCTCCCATGTAGGGCTCTCTGGTGGCCATC	1109
QY	360	SerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCys	379
DB	1110	TCCGACAGGCTCTGATGTTTCCCTCGGGCTGGGCTGGCTGGCTGGTGTAGGCAGCATGTGC	1169
QY	380	LeuPheIleAlaGlyPheAlaValGlyTyrGlyProIleProTrpLeuLeuMetSerGlu	399
DB	1170	CTCTTCATCGCTGGTGTGGAGTAGGCTGGGACCCATCCCTGGCTCTCATGTGACAGAG	1229
QY	400	IlePheProLeuHisValLysGlyValAlaIleThrGlyIleCysValLeuThrAsnTrpLeu	419
DB	1230	ATCTTCCCTGACATCAAGGGTGTGGCTACCGGGCTCTGTGCTCCACCACTGGTTC	1289
QY	420	MetAlaPheLeuValThrLysGluPheSerLeuMetGluValLeuArgProTyrGly	439
DB	1290	ATGGCCCTTCTGGTGACCAAGAGTTTAAACAGCATCTGGAGATCTCCAGCCCTACGGC	1349
QY	440	AlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal	459
DB	1350	GCCTTCGGCTCACCGGCTGCTTGTATCTCAGGCTCCTTTTCAGGCTCACCTTTGTC	1409
QY	460	ProGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg	477
DB	1410	CCTGAGACTAAGGACGAGCTCTGGAAACAATCACAGCCCTTTTCAGGGGACGG	1463
RESULT	7		
AAF55867			
ID	AAF55867	standard; cDNA; 2072 BP.	
XX	AAF55867;		
XX	17-APR-2001	(first entry)	
XX	Murine GLUTX1	coding sequence.	
XX			
KW		Murine; GLUTX; gene therapy; vaccine; hexose transport modulator;	
KW		hexose transport disorder; ischaemia; diabetes; hyperglycaemia; ss;	
XX		hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.	
OS		Mus sp.	
XX	WC200104145-A2.		
XX	18-JAN-2001.		
XX	14-JUL-2000;	2000WO-1B001042.	
XX	14-JUL-1999;	99US-0143907P.	
XX	23-AUG-1999;	99US-0151140P.	
XX	23-FEB-2000;	2000US-0184285P.	
XX	13-JUL-2000;	2000US-0061613Z.	
XX		(UWLA-) UNIV LAUSANNE.	
XX		Thorens B, Ibberson M, Uldry M;	
XX		WPI; 2001-112615/12.	
XX		P-PSDB; AAB66934.	

XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in the
 PT prevention, diagnosis and treatment of hexose transport disorders, e.g.
 PT ischemia and diabetes.
 XX Claim 3; Page 73-74; 124pp; English.
 XX

CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facultative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is the coding sequence
 CC for murine GLUTX1
 XX

SQ Sequence 2072 BP; 351 A; 673 C; 584 G; 464 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.33e-177 Length: 2072
 Score: 2150.00 Matches: 409
 Percent Similarity: 92.47% Conservative: 33
 Best Local Similarity: 85.56% Mismatches: 34
 Query Match: 87.51% Indels: 2
 DB: 4 Gaps: 2

US-09-886-954A-1 (1-477) x AAF55867 (1-2072)

QY	1	MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla	20
DB	21	ARGTCTCCAGAGACCCCGAGAGCGCAGCCGCTATTGCGGCCACCGGAAGCCAGGACT	80
QY	21	ProArgGlyArgValPheLeuAlaPheAlaAlaPheAlaAlaLeuGlyProLeuSerPhe	40
DB	81	CGCCGCGCGCGCGGCTCTCTCGCTTCTTCGCGCGCGCTGGGACCCCTCAACTTC	140
QY	41	GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro	60
DB	141	GGCTTCGCGCTCGCTACAGCTCCCGCGCATCCCGCGCAGCCCTGCGGCCACCGCCCG	200
QY	61	AlaProArgLeuAspAspAlaAlaAlaSerTyrPheGlyAlaValValThrLeuGlyAla	80
DB	201	GCCCTCGCCTCGGAGCAATGCGCCTCTGCTGGTTCGGGGCGCTGTGACCCCTGGGCGCT	260
QY	81	AlaAlaGlyValLeuGlyTyrLeuValAspArgAlaGlyArgValLeuSerLeu	100
DB	261	GCTGCAGGGGCACTACTGGCGGCTGCTCCTGGACCTTCCAGGGCGCAAGCTGAGCCTC	320
QY	101	LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal	120
DB	321	TTGCTTCGCAGCGTGCCTTCGCTGACTGGCTTGTCTCATCCCGCGCGCGGGATGTC	380
QY	121	TyrMetLeuLeuGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal	140
DB	381	TGGATGCTCTCGGAGCGCCCTCCACCGGCGCTTGCCTGGGAGTCGCCCTCCTCCTAGTG	440
QY	141	AlaProValTyrIleSerGluIleLeuTyrProAlaValArgGlyLeuLeuGlySerCys	160
DB	441	GCACCGGTTTACATCTCGAATCGCTACCCAGCTGTCGAGGACTGCTCGGCTCCTGT	500
QY	161	ValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTyrValLeuGlu	180
DB	501	GTCGAGCTGATGTTGTCATCGCATCCTCGGCTATGTCGAGCTGGTTCCTTAGAG	560
QY	181	TyrArgTyrLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys	200
DB	561	TGGCGCTGGCTGGCGTGTGGCTGTGGCCCGCCCGCCCTCATGCTGTCTGTCTGTGTC	620
QY	201	PheMetProGluThrProArgPheLeuLeuThrGlnHisArgGlnGlnAlaMetAla	220
DB	621	TACATGCCCGAGACCCAGTTTCTCTCACTCAACACAGTACCAGGAGGCGCATGGCT	680
QY	221	AlaLeuArgPheLeuTyrGlySerGluGlnGlyTyrGluAspProProIleGlyAlaGlu	240

DB	681	GCCTTCGCTCTCTGGGGCTCTGAGGAGGGCTGGAAAGAGCCCTCTGTGGGCTGAG	740
QY	241	---GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIleGlySerProPheIleIle	259
DB	741	CACCAAGGGCTTCAGCTGGCCCTGCTGAGGGCCCTGGGCATCTACAAAGCCCTCATCATC	800
QY	260	GlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAla	279
DB	801	GGCATTTCCCTCATGGCTTTCAGCAGCTGTCCAGGGTCAATGCTATCATGTTCTATGCC	860
QY	280	GluThrIlePheGluAlaLysPheLysPheLysAspSerSerLeuAlaSerValValGly	299
DB	861	AAACAGCATCTTCAGGAGGCCAAGTTCAAGGACAGCAGCCCTGGCTCGTCTGCTGGCC	920
QY	300	ValIleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArgGAG	319
DB	921	ATAATCCAGGCTCTGTCTACTGCTGTGGGGCCCTCATCATGGACAGAGCAGGGCGAAGG	980
QY	320	LeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyr	339
DB	981	CTGCTCTGGCCCTTTCGGGTGTGATCATGGTGTTCAGTATGAGTGCCTTGGTACTTAC	1040
QY	340	PheLysLeuThrGlnGlyProGlyAenSerSerHisValAlaIleSerAlaProVal	359
DB	1041	TTCAAACTGACCCAGAGCCCTCCAGCAACTCTCCCAAGTGGGCTG---GTGCCCATC	1097
QY	360	SerAlaGlnProValAspAlaSerValGlyLeuAlaTyrLeuAlaValGlySerMetCys	379
DB	1098	CGGGCGGAGCCCTGGATGTCCAAAGTGGGACTGGGCTGGCTGGTGTAGGCAGCATGTGC	1157
QY	380	LeuPheIleAlaGlyPheAlaValGlyTyrGlyProIleProTyrLeuLeuMetSerGlu	399
DB	1158	CTCTTCATTTGCTGGCTTTCGGGTGGCTGGGACCCATCCCTGGCTCTCATGTCCAGAG	1217
QY	400	IlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrIleThrLeu	419
DB	1218	APCTTCCCTCTGCAATGTCAAGGGTGTGGCTACCGGCATCTGTCTCCTCACCACCTGGTTC	1277
QY	420	MetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGly	439
DB	1278	ATGGCTTCTTAGTGACCAGAGTTCACAGGTTCAACAGGTTCAAGGATGCTCAGACCCCTACGGT	1337
QY	440	AlaPheTyrLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal	459
DB	1338	GCCTTCTGGCTCACCGCTTCTCGCTCAGTCTCAGTCTCATTCACACTGACCGTTC	1397
QY	460	ProGluThrLysGlyLysThrLeuGluIleThrAlaHisPheGluGlyArg	477
DB	1398	CCTGAGACTAAAGGCGAGGACTCTGGAAACAAGTTCACAGCCCAATTCGAGGGAGCA	1451

RESULT 8

ABX34029	ID	ABX34029 standard; cDNA; 1954 BP.
XX	AC	ABX34029;
XX	DT	10-FEB-2003 (first entry)
XX	XX	Human cancer suppressing protein PF7425.
XX	KW	Human; gene; ss; cancer suppressing protein; cancer.
OS	OS	Homo sapiens.
XX	PN	CN1351081-A.
XX	XX	29-MAY-2002.
XX	PD	31-OCT-2000; 2000CN-00127102.
XX	PF	31-OCT-2000; 2000CN-00127102.
XX	XX	

(SHAN-) SHANGHAI INST ONCOLOGY.

PA Gu J;
XX WPI; 2002-609437/66.
XX P-PSDB; ABU11283.
XX New human protein with cancer cell growth suppressing function and a
XX PT polynucleotide encoding it, for treating diseases, such as, cancer.
XX

XX Claim 5; Page 23-24 (disclosure); 39pp; Chinese.

XX This invention relates to the cDNA and protein sequences of a novel human
XX CC protein with cancer suppressing function. The invention also comprises a
XX CC method for preparing the polypeptide by recombination, and an application
XX CC of the polypeptide in treating diseases such as cancer, etc. Also
XX CC disclosed in an antagonist of the polypeptide and its medical action. The
XX CC present sequence represents a cDNA encoding a cancer suppressing protein
XX CC of the invention
XX

XX SQ Sequence 1954 BP; 325 A; 633 C; 573 G; 423 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,52e-175 Length: 1954
Score: 2123.50 Matches: 426
Percent Similarity: 89.14% Conservatave: 1
Best Local Similarity: 88.94% Mismatches: 1
Query Match: 6 6 Indels: 51
DB: Gaps: 1

US-09-886-954A-1 (1-477) x ABX34029 (1-1954)

QY 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
DB 1 GTGACGCCCGAGGACCCAGAGAAACCAGCGCTTCTGGGGCTCTCTGGGGCGCGCG 60
QY 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaAlaLeuGlyProLeuSerPhe 40
DB 61 CCGCGGGCGCGCGCGCTTCTCTGGCGCGCTTCTGGCGCGCTTCTGGCGCGCTTCTGG 120
QY 41 GlyPheAlaLeuGlyTyrosSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
DB 121 GCCTTCGGCTGGGTACAGCTCCCGGCGCATCCCTAGCTGCAGCGCGCGCGCGCGCG 180
QY 61 AlaProGluLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrIleuGlyAla 80
DB 181 GCCCGCGCGCTGGACGACGCGCGCGCTTCTGGGGCTTCTGGGGCTTCTGGGGCTTCT 240
QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgGlyLeuSerLeu 100
DB 241 CGCGCGGGGGAGTGTGGGGCGCTGGTGGACCGCGCGCGGGGCGAAGCTGAGCCTC 300
QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
DB 301 TTGCTGTGCTCCGTCCTTCTGGCGCGCTTCTGGCGCGCTTCTGGCGCGCTTCTGG 360
QY 121 TrpMetLeuLeuGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaAlaSerLeuVal 140
DB 361 TGGATGTGCTGGGGGGCGCGCTTCTCACCGCGCGCTTCTGGCGCGCTTCTGGCGCG 420
QY 141 AlaProValTyrosSerGluIleAlaValAlaValAlaValAlaValAlaValAlaVal 160
DB 421 GCCCGCGCTTACATCTCGAATTCGCTTACCAGCAGTCCCGGGGTGGTGGCTCCTG 480
QY 160 sValGlnLeuMetValValValGlyIleLeuLeuAlaTyrosAlaGlyTrpValLeuG 180
DB 481 TGTGAGCTTAATGT 540
QY 180 uTrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
DB 541 GTGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 600
QY 200 sPheMetProGluThrProArgPheLeuLeuThrGlnIleHisArgGlnGluAlaMetAl 220

601 CTTTCATGCCCCGAGACCCCGCGCTTCTGCTGACTCAGCACAGGCGCCAGGAGCCATGGC 660
QY 220 aAlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGl 240
DB 661 CGCCCTGGCGTTCCTGTGGGGCTCCGAGCAGGGCTGGGAAGACCCCGCCATCGGGCTGA 720
QY 240 uGlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrosLeuProPheIleIleG 260
DB 721 GCAGAGCTTTCACCTGGCGCTTCTGGCGAGCCCGGCACTTCAAGCCCTTCATCATCGG 780
QY 260 YValSerLeuMetAlaPheGlnGlnLeuSerGlyValAenAlaValMetPheTyroAlaGl 280
DB 781 CGTCTCCCTGATGGCTTCCAGCAGCTGTCCGGGGTCAACGCGCTCATGTTTATGCGAGA 840
QY 280 uThrIlePheGluAlaIlePheIlePheIlePheIlePheIlePheIlePheIlePhe 300
DB 841 GACCACTTTTGAAGAGGCCAAGTTCAAGGACAGCAGCCCTGGCGCTGGCTGGCTGG 900
QY 300 lIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgLe 320
DB 901 CATCCAGGTGCTGTTTCAAGCTGTGGCGCTTCTCATCATGGACAGAGCGCGAGGCT 960
QY 320 uLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTrpPh 340
DB 961 GCTCCTGCTTCTGAGGTGTGGTTCATGTGTGTTTACGACAGGTGCTTCCGGCGCTTCT 1020
QY 340 eLysLeuThrGlnGlyGlyProGlyAsnSerHisValAlaIleSerAlaProValSe 360
DB 1021 CAAGCTGACCCGAGGGTGGCCCTGGCAACTCTCCGACGTGGCCATCTCGGCGCTGCTC 1080
QY 360 rAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLe 380
DB 1081 TGCACACCTTGTGATGTCAGCTGGGGCTGGCGCTGGCTGGCTGGCGCGAGCATGTCCT 1140
QY 380 uPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluI 400
DB 1141 CTTTCATCGCGGAGGTC----- 1157
QY 400 lPheProLeuHisValValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuM 420
DB 1157 ----- 1157
QY 420 eAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTrpGlyVA 440
DB 1158 -----CTCAGGCCCTATGGAG 1173
QY 440 lAPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValP 460
DB 1174 CTTTCTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 1233
QY 460 roGluThrLysGlyLysThrLeuGlnIleThrAlaHisPheGluGlyArg 477
DB 1234 CTGAACTAAAGAAAGACTCTGGAAACAAATCACAGCCCATTTTGGGGGGCGGA 1286
RESULT 9
ACN43514
ID ACN43514 standard; cDNA; 1923 BP.
XX AC ACN43514;
XX XX 18-NOV-2004 (first entry)
DB XX Human diagnostic and therapeutic polynucleotide SEQ ID NO:2389.
XX ss; gene; gene therapy; human diagnostic and therapeutic polynucleotide;
XX ditup.
XX OS Homo sapiens.
XX XX WO2004023973-A2.
XX XX 25-MAR-2004.
XX PD

Db 943 ctctggtctgtcagctgggtcanggtgttcagcagcagctgcttccggcctacttc 1002
 QY 341 LvsLeuThrGlnGlyProGlyAemSerHisValAlaIleSerAlaProValSer 360
 Db 1003 AAGCTGACCCAGGGTGGCCCTGGCAACTCTCGCACGTGGCCATCTCGGCCCTGTCTCT 1062
 QY 361 AlaGlnProValAspAlaSerValGlyLeuAlaThrLeuAlaValGlySerMetCysLeu 380
 Db 1063 GCACAGCCTTGTATGCCAGCGTGGGCTGGCCCTGGCTGGCCGCTGGCGACATGTGCTC 1122
 QY 381 PheIlealaGlyPheAlaValGlyTrpGlyProIleProIleProIleProIleLeu 400
 Db 1123 TTCATGCCCGCTTGGCGTGGGCTGGGCGCCATCCCTGGTCTCAATGTTCAGAGATC 1182
 QY 401 PheProLeuHisValGlyValAlaThrGlyLeuValGlyLeuValLeuThrLeuMet 420
 Db 1183 TTCCTCTGATGTCAAGGGCTGGCGACAGGATCTGGCTCTCAACCACTGGCTCATG 1242
 QY 421 AlaPheLeuValThrIleGlyPheSerSerLeuMetGluValLeuArgProIleGlyAla 440
 Db 1243 GCCTTCTCGTACCAAGAGTTCAGCAGCCTCATGGAGGTCTTCAGGCCCTATGGAGCC 1302
 QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
 Db 1303 TTCGGCTTGCCTCCGCTTTCGATCTTCAGTGTCTCTTTTTCATTTTGTGTCCCT 1362
 QY 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 Db 1363 GAACTAAGGAAGACTCTGGAAACAATCACAGCCCAATTTTGGGGGCGA 1413

RESULT 11
 ADL33401
 ID ADL33401 standard; DNA; 1790 BP.
 XX ADL33401;
 AC XX
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Human transporter and ion channel (TRICH) gene #46.
 XX
 KW anti-HIV; anti-allergic; anti-inflammatory; antianemic; antiparkinsonian;
 KW neurotropic; anticonvulsant; antiarteriosclerotic; antiasthmatic;
 KW immunosuppressive; antihypertensive; antidiabetic; hepatotropic; dermatological;
 KW antidiabetic; nephrotropic; antitumor; thyromimetic; neuroprotective;
 KW osteopathic; antiarthritic; antiparasitic; antihelminthic; antipsoriatic;
 KW uropathic; ophthalmological; antirheumatic; haemostatic; antibacterial;
 KW virucide; protozoacide; fungicide; gene therapy; gene; ds.
 XX
 OS Homo sapiens.
 XX

XX 09-OCT-2003.
 XX 27-MAR-2003; 2003WO-US009797.
 XX 28-MAR-2002; 2002US-0368840P.
 XX 26-APR-2002; 2002US-0375637P.
 XX (INCY-) INCYTE CORP.
 XX Marquis JP, Lee SY, Emerling BM, Hafalia AJA, Khare R, Kable AE;
 XX Richardson TW, Swarnakar A, Chawla NK, Becha SD, Mason PM;
 XX Elliott VS, Ramkumar J, Griffin JA, Tran UK, Ison CH, Lindquist EA;
 XX Jiang X, Jackson AA, Wilson AD, Jin P, Chang H;
 XX WPI: 2003-833535/77.
 XX P-PSDB; ADL33342.

PT New human transporters and ion channels (TRICH) and polynucleotides,
 PT useful for diagnosing, treating or preventing autoimmune or inflammatory
 PT disorders (e.g. AIDS, allergy or anemia), multiple sclerosis, cancer or
 PT hepatitis.
 XX
 PS Claim 12; SEQ ID NO 105; 405pp; English.
 XX
 CC The invention relates to an isolated polypeptide (I), which is a human
 CC intracellular signaling molecule, which is a human intracellular
 CC signaling molecule, a naturally occurring amino acid sequence at least 90%
 CC identical to it or a biologically active fragment or an immunogenic
 CC fragment of the polypeptide. The human TRICH, polynucleotides, agonists
 CC and antagonists are useful for diagnosing, treating or preventing
 CC disorders associated with aberrant expression of TRICH, particularly cell
 CC proliferative disorders (e.g. arteriosclerosis, atherosclerosis,
 CC cirrhosis, hepatitis, paroxysmal nocturnal hemoglobinuria, polycythemia
 CC vera, psoriasis, primary thrombocytopenia or cancer), developmental
 CC disorders (e.g. renal tubular acidosis, anemia or mental retardation),
 CC neurological disorders (e.g. Alzheimers disease, Parkinson's disease or
 CC epilepsy), autoimmune/inflammatory disorders (e.g. AIDS, allergies,
 CC asthma, autoimmune thyroiditis, contact dermatitis, Crohn's disease,
 CC diabetes mellitus, glomerulonephritis, Goodpasture's syndrome, gout,
 CC Graves' disease, Hashimoto's thyroiditis, irritable bowel syndrome,
 CC multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis, Reiter's
 CC syndrome, rheumatoid arthritis, Sjogren's syndrome, uveitis), or viral,
 CC bacterial, fungal, parasitic, protozoan or helminthic infections. The
 CC polynucleotides encoding TRICH are useful for creating transgenic animals
 CC to model human disease. This sequence corresponds to the gene encoding
 CC one of the proteins of the invention.
 XX

XX SQ Sequence 1790 BP; 279 A; 587 C; 535 G; 389 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 3 42e-149 Length: 1790
 Score: 1823.50 Matches: 370
 Percent Similarity: 77.41% Conservative: 0
 Best Local Similarity: 77.41% Mismatches: 0
 Query Match: 74.22% Indels: 108
 DB: 11 Gaps: 1
 US-09-886-954A-1 (1-477) x ADL33401 (1-1790)
 QY 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProGlyGlySerAla 20
 Db 47 ATGACGCGCCGAGGACCCAGAGAAACCCAGCCGCTTCGGGGCTCTCTGGCGGCGCG 106
 QY 21 ProArgGlyArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
 Db 107 CCCTGGCGCGCGCGCTTCCTCGCCGCTTCGCGCTTCGCGCTTCGCGCTTC 166
 QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaProPro 60
 Db 167 GCTTCGCGCTCGCTACAGTCCCGCCATCCCTAGCTTCGAGCGCGCGCGCGCG 226
 QY 61 AlaProArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValThrLeuGlyAla 80

CC protozoal (or fungal), metabolic disorders (e.g. obesity), reproductive
 CC disorders (e.g. infertility), neurological disorders (Parkinson's
 CC disease, Alzheimer's disease, stroke, epilepsy), cardiovascular disorders
 CC (e.g. myocardial infarction or hypertension), eye disorders, or cell
 CC proliferative diseases (e.g. cancer, atherosclerosis, hepatitis). These
 CC are also useful in assessing the effects of exogenous compounds on the
 CC expression of nucleic acid and amino acid sequences of MDR1. The MDR1 or
 CC its fragments are useful in screening compounds for effectiveness as
 CC agonist or antagonist of the polypeptides, or in altering the expression
 CC of the target polynucleotide and compounds that specifically bind to or
 CC modulate the activity of the polypeptide. The microarray of the invention
 CC is useful in monitoring or measuring protein-protein interactions, drug-
 CC target interactions, and gene expression profiles. This polynucleotide
 CC sequence represents the DNA of an MDR1 polypeptide of the invention.

SQ Sequence 1763 BP; 305 A; 543 C; 513 G; 402 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.76e-142 Length: 1763
 Score: 1746.50 Matches: 353
 Percent Similarity: 74.00% Conservative: 0
 Best Local Similarity: 74.00% Mismatches: 1
 Query Match: 71.08% Indels: 124
 DB: 12 Gaps: 1

US-09-886-954A-1 (1-477) x ADH45466 (1-1763)

QY	1	MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlySerAla	20
DB	63	ATGACCCCGAGACCAGGANAACCCAGCCGCTTCGGGGCTCTCGGGCCAG	118
QY	21	ProArgGlyArgValPheLeuAlaPheAlaAlaLeuGlyProLeuSerPhe	40
DB	118	-----	118
QY	41	GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaProPro	60
DB	118	-----	118
QY	61	AlaProArgLeuAspAlaAlaSerTropheGlyAlaValThrLeuGlyAla	80
DB	118	-----	118
QY	81	AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu	100
DB	118	-----	118
QY	101	LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal	120
DB	118	-----	118
QY	121	TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal	140
DB	118	-----	118
QY	141	AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys	160
DB	119	-----GTCTACATCTCCGAAATCGCTACCCAGCAGTCCGGGGTTCCTCGGCCTCT	172
QY	161	ValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu	180
DB	173	GTGACGTAATGTCGTCGCGCATCTCTGGCCCTACCTGGCAGCTGGGTGGTGGAG	232
QY	181	TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuMetCys	200
DB	233	TGGCGTGGCTGCTGCTGGGCTGGGCTGGCCCTCCCTCATGCTGCTTCATGCTG	292
QY	201	PheMetProGluThrProArgPheLeuLeuThrGlnHisArgGlnGluAlaMetAla	220
DB	293	TTTCATGCCGAGACCCCGGCTTCCTGCTGACTCAGCAGAGCCAGGAGCCATGGCC	352
QY	221	AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProPheLeuAlaGlu	240

DB	353	GCCTCGCGTTCCTCTGTTGGGGCTCTGACGAGGGGCTGGGAAGACCCCCCTCGGGGCTGAG	412
QY	241	GlnSerPheHisLeuAlaLeuArgGlnProGlyIleTyrIleTyrIleGly	260
DB	413	CAGAGCTTTCACCTGGCCCTGCTGGGAGCCCGGCACTTACAAAGCCCTTCATCATCGGC	472
QY	261	ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu	280
DB	473	GTCTCCCTGATGGCTTCACGAGCTGTCGGGGGTCAACCGCTCATGTTCTATGACGAG	532
QY	281	ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVal	300
DB	533	ACCATCTTTTGAAGAGGCCAAAGTTCAAGGACAGCAGCCCTGGCTCGTGGGTGTC	592
QY	301	IleGlnValLeuPheThrAlaValAlaLeuLeuMetAspArgAlaGlyArgArgLeu	320
DB	593	ATCCAGGTGCTGTTTACAGCTGTGGCGGCTCTCATGTGACAGACAGAGCCGGGAGGCTG	652
QY	321	LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe	340
DB	653	CTCTCTGGTCTTCTCAGGTGCTGCTCATGTGTTCAGCACGAGTGCCTTCGGCGCTACTTC	712
QY	341	LysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer	360
DB	713	AACTGACCCAGGGTGGCCCTGGCAACTCTCTCGACGTCGCCATCTCGGGCGCTGCTCT	772
QY	361	AlaGlnProValAspAlaSerValGlyLeuAlaTyrLeuAlaValGlySerMetCysLeu	380
DB	773	GCACAGCCTTGTGATGCCAGCGTGGGCTGGCTGGCTGGCGGAGCAGCATGTGCCCTC	832
QY	381	PheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIle	400
DB	833	TTCTATCGCCGGCTTTCGGGTGGGTCGGGGCCCATCCCTGCTCTCATGTCCAGAGATC	892
QY	401	PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet	420
DB	893	TTCCCTCTGCATGTCAAGGGCGTGGCGACAGGCATCTGCGTCTCCACCAACTGGCTCATG	952
QY	421	AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla	440
DB	953	GCCTTTCGTGACCAAGAGTTCAGCAGCCTTCATGGAGGTCTCAGGCCCTATGGGCC	1012
QY	441	PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro	460
DB	1013	TTCTGGCTTTCCTCCGCTTTCGTCATCTTCAGTGTCTTTCACCTTGTCTGTGTCCT	1072
QY	461	GluThrLysGlyThrLeuGluGlnIleThrAlaHisPheGluGlyArg	477
DB	1073	GAACCTAAAGGAAGACTCTGGAAACAAATCACAGCCCAATTTTGAGGGGGCA	1123
RESULT	13		
ABL90126			
ID	ABL90126	standard; cDNA; 1156 BP.	
AC	ABL90126;		
DT	24-MAY-2002	(first entry)	
XX			
DE	Human polynucleotide SEQ ID NO 688.		
XX			
KW	Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;		
KW	antiallergic; hepatotropic; antidiabetic; anti-inflammatory; antiulcer;		
KW	vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;		
KW	cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;		
XX	neurological disease; infection; human; secreted protein; gene; ss.		
OS	Homo sapiens.		
XX			
PN	WO200190304-A2.		
XX			
PD	29-NOV-2001.		
XX			
PF	18-MAY-2001; 2001WO-US016450.		

XX AAD12574;
 AC XX
 DT DT
 XX XX
 DE DE
 XX XX
 KW Human protein having hydrophobic domain encoding cDNA clone HP10784.
 KW Human; hydrophobic domain; gene therapy; nutritional supplement;
 KW cell proliferation; immunomodulatory; autoimmune disorder; antimicrobial;
 KW multiple sclerosis; rheumatoid arthritis; insulin-dependent diabetes;
 KW haematopoiesis; tissue growth activity; Parkinson's disease; cytostatic;
 KW Huntington's disease; Alzheimer's disease; chemotactic; chemokinetic;
 KW haemostatic; thrombolytic; tumour growth inhibitor; anabolic;
 KW contraceptive; antiinfertility; antiinflammatory; ss.
 XX OS
 XX XX
 XX XX
 FH Location/Qualifiers
 FT 61.849
 FT CDS
 FT /tag= a
 FT /product= "Human protein having hydrophobic domain"
 FT /note= "CDS is specifically is claimed in claim 3"
 XX
 PN W0200149728-A2.
 XX
 XX
 PD 12-JUL-2001.
 XX
 PD 28-DEC-2000; 2000WO-JP009359.
 PF
 XX
 XX
 PR 06-JAN-2000; 2000JP-0000585.
 PR 06-JAN-2000; 2000JP-0000588.
 PR 11-JAN-2000; 2000JP-00002299.
 PR 03-FEB-2000; 2000JP-00026862.
 PR 03-MAR-2000; 2000JP-00059367.
 XX
 XX (PROT-) PROTEGENE INC.
 PA (SAGA) SAGAMI CHEM RES CENT.
 XX
 XX Kato S, Kimura T;
 DR WPI: 2001-418355/44.
 DR P-FSDB; AAS06579.
 XX
 XX Human proteins with hydrophobic domains and the nucleic acids encoding
 PT them, useful for preventing diagnosing and treating e.g. cancer,
 PT Alzheimer's and inflammation.
 XX
 PS Claim 4; Page 287-289; 563pp; English.
 XX
 CC The present sequence is human protein with hydrophobic domain encoding
 CC cDNA clone HP10784. The polynucleotide and polypeptide of the invention
 CC may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate polypeptide expression. The polynucleotides
 CC may be used to produce the polypeptide, by inserting the nucleic acids
 CC into a host cell and culturing the cell to express the protein. The
 CC polynucleotides and its complementary sequences may also be used as DNA
 CC probes in diagnostic assays and also used in gene therapy. The
 CC polypeptides may also be used as antigens in the production of antibodies
 CC and in assays to identify modulators of polypeptide expression and
 CC activity. The polypeptides and nucleic acids may be used as nutritional
 CC supplements, to modulate cytokine and cell proliferation activity, to
 CC modulate immune stimulation or suppression (e.g. for the treatment of
 CC microbial infections and autoimmune disorders such as multiple sclerosis,
 CC rheumatoid arthritis and insulin-dependent diabetes), to modulate
 CC haematopoiesis, to modulate tissue growth activity (e.g. for the
 CC treatment of Parkinson's disease, Huntington's disease and Alzheimer's
 CC disease), to modulate activin and inhibin activity (e.g. for controlling
 CC fertility), to modulate chemotactic and chemokinetic activity, to
 CC modulate haemostatic and thrombolytic activity, to modulate receptor
 CC ligand activity, to modulate inflammation and to inhibit tumour growth
 XX
 XX Sequence 1461 BP; 195 A; 501 C; 456 G; 309 T; 0 U; 0 Other;

Alignment Scores: 7.84e-126 Length: 1461
 Pred. No.: 1554.50 Matches: 324
 Score: 67.85% Conservative: 1
 Best Local Similarity: 67.64% Mismatches: 1
 Query Match: 63.27% Indels: 153
 DB: Gaps: 2
 US-09-886-954A-1 (1-477) x AAD12574 (1-1461)
 QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
 Db 61 ATGACGCCCGAGGACCCAGAGAAACCCAGCGGCTTCTGGGCGCTCTCTGGCGGACGCGG 120
 QY 21 ProArgGlyArgValPheLeuAlaAlaPheAlaAlaLeuAlaLeuGlyProLeuSerPhe 40
 Db 121 CCCCAGCGCGCGCGGCTTCTCTCGCGGCTTCTCGCGGCTTCTGGGCGCTTCTGGGCGGCTTCT 180
 QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 Db 181 GGCCTTCGGCTCGGCTACAGCTCCCGCGGCATCCCTAGCTGCAGCGCGCGCGCGCGCGCG 240
 QY 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
 Db 241 GCCCGCGCGCTGGAGCGCGCGCGCTCTCTGGTTCGGGCGCTCTCTGGGCGCTTCTGGGCGG 300
 QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgGlySerLeu 100
 Db 301 GCGCGCGCGGAGTGTGGCGGCTGGCTGGTGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
 QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 Db 361 TTGCTGTCTCGTCCGCTCCCTCTGGCGGCTTCTGGGCGCTTCTGGGCGCTTCTGGGCGG 420
 QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 Db 421 TGGATGTCTGGGCGCGCGCTCTCTCACCGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGG 480
 QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
 Db 481 GCCCGGCTCACTCTCGAAATCGCTTACCAGCAGTCCGGGGGTTGTCTGGCTCTCTGGCTCT 540
 QY 161 ValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
 Db 541 GTGAGCTAAATGCTCGTGGCGCATCTCTGGCGCTTCTGGGCGCTTCTGGGCGCTTCTGG 600
 QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuLeuMetCys 200
 Db 601 TGGCGCTGGCTGGCTGGCTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGG 660
 QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAla-MetAl 220
 Db 661 TTCATGCCCGAGACCCCGCGCTTCTGTGACTCAGCAGCGCGCGCGCGCGCGCGCGCGCGCG 718
 QY 220 eAlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaG 240
 Db 718 ----- 718
 QY 240 uGlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIleG 260
 Db 718 ----- 718
 QY 260 yValSerLeuMetAlaPheGlnLeuSerClyValAaAlaValMetPheTyrAlaG 280
 Db 718 ----- 718
 QY 280 uThrIlePheGluAlaAlaLysPheLysAspSerSerLeuAlaSerValValGlyVa 300
 Db 718 ----- 718
 QY 300 IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgGlu 320
 Db 718 ----- 718

CC The present invention relates to GLUTX proteins (AAF55871 and AAF55872-954A-1). The GLUTX proteins are related to the facultative glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport function. The GLUTX proteins may be used in the diagnosis, prevention and treatment of hexose transport disorders such as ischaemia, diabetes, hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a neurodegenerative disease. The present sequence is the coding sequence for rat GLUTX3

SQ Sequence 2011 BP; 337 A; 621 C; 593 G; 459 T; 0 U; 1 Other;

Alignment Scores: 7, 47e-75 Length: 2011
 Pred. No.: 972.50 Matches: 219
 Percent Similarity: 61.19% Conservative: 79
 Best Local Similarity: 44.97% Mismatches: 169
 Query Match: 39.58% Indels: 21
 Gaps: 7

CC AAB66932-AAB66941). The GLUTX proteins are related to the facultative glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport function. The GLUTX proteins may be used in the diagnosis, prevention and treatment of hexose transport disorders such as ischaemia, diabetes, hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a neurodegenerative disease. The present sequence is the coding sequence for rat GLUTX3

US-09-886-954A-1 (1-477) x AAF55871 (1-2011)

Qy 3 ProGluAspProGluGlnThrGlnProLeuLeuGlyProProGlyGlySerAlaProArg 22
 Db 64 CCGCGCTGGCTGAAGAG-----AAGGCACGGCGGGGGCTTGC 105

Qy 23 GlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPhe 42
 Db 106 AACAGAGGGTGTTCCTGGCCACCTTCGTCGCGTAAAGCGCTTCTGACCCAGCACTA 165

Qy 43 AlaLeuGlyTyrSerProAlaIleProSerLeuGlnArgAlaAlaProAlaPro 62
 Db 166 GCCCTGTGTACACATCTCCAGTCACTCCCTGCGTAAAGCGCTTCTGACCCAGCACTA 225

Qy 63 ArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValThrLeuGlyAlaAlaAla 82
 Db 226 CGCCTGGACAAAATCCAGGCATCTGTGTGGTGGTGTTCACCTCCCTGGTGGCGCGCT 285

Qy 83 GlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeu 102
 Db 286 GGGGCGCTCAGTGTACTCAATGACTCTTGGCGCGAAGCTCAGCATCATGTTT 345

Qy 103 CysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMet 122
 Db 346 TCCGCTGTCCCTCGGCCAATGGTATGGTATGCACTATGCTGGTGGTGGCGCTGTGATG 405

Qy 123 LeuLeuGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaPro 142
 Db 406 CTTCTGTGGGGAGATGCTGACAGGCTTTCGGGGGAGCTCCTGCTGCTGCTGCTGCTGCTG 465

Qy 143 ValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGln 162
 Db 466 GTGTACGTGTCTGAGATTGACCCCGGGTGTTCGGCGGGCCCTGGGGGGCCACCGCCGAG 525

Qy 163 LeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGluTrpArg 182
 Db 526 CTCATGGCCGTGTTGGATCCCTCTCTATGCCCTGTGTCTCTGCTGCTGCTGCTGCTGCTG 585

Qy 183 TrpLeuAlaValLeuGlyCysValProProSerLeuLeuLeuLeuMetCysPheMet 202
 Db 586 TGGCTTGGCTGGCGGGAGGGGCTGTTCTGGTCAATGATCTCTGCTGCTGCTGCTGCTGCTG 645

Qy 203 ProGluThrProArgPheLeuLeuThrGlnHisArgGlnGluAlaMetAlaAlaLeu 222
 Db 646 CCCAACTGGCTGCTTCTAATGAGCCGGATGAGGAGCACTGAGCCGCGCTG 705

Qy 223 ArgPheLeuTrp-----GlySerGluGlnGlyTrpGluAspProPheIleGly--- 238
 Db 706 -----AICTGGCTCGAGCGGACTCTGAGGTCCACTGGAGTGTGAGCAGATCCAGGAC 759

Qy 239 -----AlaGluGlnSerPheHisLeuAlaLeuLeuArg-----GlnProGlyIleTyr 254
 Db 760 AACGTGGGAGACAGAGTAGCCGAGTGTCTGGTGGGGGAGCGCTGGGAGCCCGCTGTGTAC 819

RESULT 15
 AAF55871 standard; cDNA; 2011 BP.

ID AAF55871
 AC AAF55871;
 DT 17-APR-2001 (first entry)
 XX Rat GLUTX3 coding sequence.
 DE Rat GLUTX3 coding sequence.
 XX Rat; GLUTX3; gene therapy; vaccine; hexose transport modulator;
 KW hexose transport disorder; ischaemia; diabetes; hypoglycaemia; as;
 KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 XX Rattus sp.
 OS WO200104145-A2.
 PN 18-JAN-2001.
 XX 14-JUL-2000; 2000WO-IB001042.
 XX 14-JUL-1999; 99US-0143907P.
 PR 27-AUG-1999; 99US-0151140P.
 PR 23-FEB-2000; 2000US-0184285P.
 PR 13-JUL-2000; 2000US-00616132.
 XX (UYLA-) UNIV LAUSANNE.
 XX Thorens B, Ibberson M, Uldry M;
 DR WPI; 2001-112615/12.
 DR P-ESDB; AAB66938.
 XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in the
 PT prevention, diagnosis and treatment of hexose transport disorders, e.g.
 PT ischaemia and diabetes.
 XX Claim 3; Page 82-83; 124pp; English.

Qy 320 uLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPh 340
 Db 719 ---CTGTCCTGTTCAGGTGGTTCATGGTGTTCAGCACGAGTGCCTTGGCGGCTACTT 774

Qy 340 eLysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSe 360
 Db 775 CAAGCTGACCCAGGGTGGCCCTGGCAACTCTCCGACGCTGGCCATCTCGGGCGCTGCTC 834

Qy 360 rAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLe 380
 Db 835 TGCACAGCCTGTGTATGCCAGCGTGGCGCTGGCTGGCTGGCGGACGATGTGGCT 894

Qy 380 uPheIleAlaGly-PheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluI 400
 Db 895 CTTCACTGCCGGAGGTC----- 911

Qy 400 lePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuM 420
 Db 911 ----- 911

Qy 420 etAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyA 440
 Db 912 -----CTCAGGCCCTATGGAG 927

Qy 440 laPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValP 460
 Db 928 CCTTCTGGCTTGCCTCCGCTTCTGCATCTTCAGTGTCTTCACTTTGCTTGTCTGTGCC 987

Qy 460 roGluThrLysGlyThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 Db 988 CTGAATAAGAAAGACTTGGAAACAATAACACGCCCAATTTTGGGGGGCGA 1040

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 25, 2005, 18:36:27 ; Search time 236 Seconds
(without alignments)
3307.220 Million cell updates/sec

Title: US-09-886-954A-1

Perfect score: 2457
Sequence: 1 MTPDPPEPQPLPPGSSA.....CVPETKGTLEQITAFHEGR 477

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DRV=xlp
-O=/cgn2_1/USPTO_spool_p/US09886954/runat_23022005_154243_13234/app_query.fasta_1.647
-DB=Issued Patents NA -QFMT=fastp -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09886954 @CGN 1.1 105 @runat_23022005_154243_13234 -NCPUS=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAFOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/1/ina/FACTUS COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Table with columns: Result No., Score, Match, Length, DB ID, Description. Contains 12 rows of search results.

Table with columns: Hit No., Score, Match, Length, DB ID, Description. Contains 45 rows of search results.

ALIGNMENTS

RESULT 1
US-09-949-016-3739 Application US/09949016
Sequence 3739 WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949, 016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3739
LENGTH: 740
TYPE: DNA
ORGANISM: Human
US-09-949-016-3739

Alignment Scores:
Pred. No.: 3,298,119 Length: 740
Score: 1215.50 Matches: 241
Percent Similarity: 99.59% Conservative: 0
Best Local Similarity: 99.59% Mismatches: 0
Query Match: 49.47% Indels: 1
DB: 4 Gaps: 1
US-09-886-954A-1 (1-477) x US-09-949-016-3739 (1-740)

```

; ORGANISM: Human
US-09-949-016-15481

Alignment Scores:      Length: 10143
Pred. No.:            Matches: 429
Score:                870.50
Percent Similarity:   15.43%
Best Local Similarity: 15.43%
Query Match:         2351
Indels:              3
Gaps:                8
DB:

US-09-886-954A-1 (1-477) x US-09-949-016-15481 (1-10143)
QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlySer---- 19
Db 1291 ATGACGCCGAGGACCAGAGGNAACCCAGCCGCTTCGGGGCCCTCTGGCGGCAG-GTG 1349
QY 19 ----- 19
Db 1350 AGCTCCGGGGAAACCCGGGCGGATGCGGCTCCGCCCTCCCGCCCTCCGCGFCGCGCTCC 1409
QY 20 -----AlaProArgGlyArgValPheLeuAlaAla 30
Db 1410 GCTCACCCCTCGGCCCTGTCCTCCCGAGCGGCGCCCGCGCGCTTCCTCGCGCC 1469
QY 31 PheAlaAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeuGlyTyrSerSerProAla 50
Db 1470 TTGGCGCGTCCCTGGGCCACTCAGCTTCGGCTTCGGCTTCAGCTCCCGCGCC 1529
QY 51 IleProSerLeuGlnArgAlaAlaProProAlaProArgLeuAspAlaAlaAlaSer 70
Db 1530 ATCCCTAGCTGACGCGCGCCCGCCCGCCCGCCCGCCCGCTGAGACAGCCGCGCCCTCC 1589
QY 71 TrpPheGly----- 73
Db 1590 TGGTTCGGGTGAGGCCCGCGGCTCGCTCCCGCCCTGGGACCCACGCGCCCTCCTCTC 1649
QY 73 ----- 73
Db 1650 GGGACGGCATCGGGACCCCTCCGCCCCCACCCTTCCCTCGGGACAGGCAATCGGGCCCC 1709
QY 73 ----- 73
Db 1710 GCCGCCACCACTTTCCTCCAGAGAGAGGCAATGGGCTCTCTGCCCCCATCCCTTCCC 1769
QY 73 ----- 73
Db 1770 TCGGGACGGGCAATCGAGCCCTCAGCCCCACCCCTCCCTCGGGACAGCACCGGGCCC 1829
QY 73 ----- 73
Db 1830 CTCAGCCCCCTACCCCTTCCCTCGGGACGGGGCGGGGGCCCGGGCTTTTACCTCTGAG 1889
QY 73 ----- 73
Db 1890 CCCGGGGAGTGGGTGGAGGGGGAGGATGGGCTTCGGGCTTCGGGCGCCCTCCTCAGCA 1949
QY 74 -----AlaValThrLeuGlyAlaAlaAlaGlyGlyValLeuGlyGly 88
Db 1950 GCCGCCCGCTCCAGCTGTGTGACCTCGCGGTCGCGGGGGGGAGTCTCGGGCGCC 2009
QY 89 TrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeuLeuLeuLeuLeuLeuLeuLeu 108
Db 2010 TGGCTGGTGGAGCCCGCGCGGCGCAAGCTGAGCCCTTGTGCTGTCTCGTCCCTTCTGTG 2069
QY 109 AlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMetLeuLeuGlyGlyArgLeu 128
Db 2070 GCCGGCTTTCGGCTCATCACCGCGCCAGAGACGTGTGGATGCTGTCTGGGGGGCGCCCTC 2129
QY 129 LeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaPro----- 142
Db 2130 CTCACCGGCTGGCTGGGTTGCTCCCTCCCTAGTGGCCCGGGTGGAGTCCCGCTCTC 2189

```

```

1 CTGGCGCGTGGCTGGTGGAGCCGGCGCGGCGCAAGCTGAGCCCTCTTGGCTGTGCTCCGCTG 60
106 PropheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMetLeuLeuGly 125
61 CCCTTGCTGGCGGGCTTGGCCGTCATCACCGCGGCGCCAGGACGTTGTGATGCTGCTGGGG 120
126 GlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaProValTyrIle 145
121 GCGCCGCTTCTCCACCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 180
146 SerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGlnLeuMetVal 165
181 TCCGAAATCCCTACCCAGCAGTCCGGGGTGGCTCGGCTCCCTGTGTGTCAGCTAATGGTC 240
166 ValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGluTrpAigTrpLeuAla 185
241 GTCGTGGGCAATCCTCTGGCCCTACCTGGCAGGCTGGGTGGGTGGGTGGGTGGGTGGCT 300
186 ValLeuGlyCysValProProSerLeuMetLeuLeuLeuMetCysPheMetProGluThr 205
301 GTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 360
206 ProArgPheLeuLeuThrGlnHisArgGlnGluAlaMetAlaAlaLeuArgPheLeu 225
361 CCGCGCTTCTGCTACTCAGCACAGCCCGCAGGAGCCATGGCCCGCCCTCGCGTTCCTG 420
226 TrpGlySerGluGlnGlyTrpGluAspProPheIleGlyAlaGlu---GlnSerPheHis 244
421 TGGGGCTCCGAGCAGGGCTGGAGAGACCCCTCCATCGGGCTGAGCAGCAGAGCTTTCAC 480
245 LeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIleGlyValSerLeuMet 264
481 CTGGCCCTGTCGGGACCGCCGCACTACAAGCCCTTCAATCGGGCTTCCCTGATG 540
265 AlaPheGlnLeuSerGlyValAanAlaValMetPheTyrAlaGlnThrIlePheGlu 284
541 GCCTTCAGCAGCTGTGCGGGGTCAACGGCGTCAATGTTCTATGCAGAGACCACTTTGAA 600
285 GluAlaLysPheLysAspSerSerLeuAlaSerValValGlyValIleGlnValLeu 304
601 GAGGCCAAGTTCAAGGACAGCAGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 660
305 PheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgLeuLeuValLeu 324
661 TTCACAGCTGGCGGCTCTCATCATGACAGAGCAGGCGGAGGCTGCTCTGGTCTTG 720
325 SerGly 326
721 TCAGGT 726

RESULT 2
US-09-949-016-15481
; Sequence 15481, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15481
; LENGTH: 10143
; TYPE: DNA

```

QY 142 ----- 142
 Db 2190 GAGTGTCTCTCGGGCCTGAGACCGAGGGGAGTGGACAAACCGTCCCCAGGCCT 2249
 QY 142 ----- 142
 Db 2250 GGGGGGGGCTCCCTCGGGGACCTCTGCGGTCAGGCTTGAAGTCCCTGCGCTTAT 2309
 QY 142 ----- 142
 Db 2310 CTGCGGTCCCTCCCGTCGACCCCTGGAGGATCCTACTGTTCTCTCCATTTTACACTGA 2369
 QY 142 ----- 142
 Db 2370 GETCATGACATGAGTCTCGMAAGGTGAAGTCCCTTTGGCCAGCGAGTCCACAGCTAG 2429
 QY 142 ----- 142
 Db 2430 TGAGGGGAAGCAGTTGCAGGAACCCAGGGTTGCCACTTAGCGTCCCGCTCTTTTG 2489
 QY 142 ----- 142
 Db 2490 CTCTGCAACCTCGGATGATCCACAGAGCCACTCCCTACATTTTGGTTTTTCATCCCT 2549
 QY 142 ----- 142
 Db 2550 GGCTTCGGGTCAATGACTGCAATTAGCAGGAGTTCCTGTCTGATTTCTCTTGGGGCAA 2609
 QY 142 ----- 142
 Db 2610 AGGCTCGGCTCTGTGTTCCCTTCCCTTCTCACAGGCCAGCCCTCTGGCTTCCCA 2669
 QY 142 ----- 142
 Db 2670 TCTTCCCTTTCCCTCTGTACCAGGCTGCCACTTGGGAGCCCGCAACATCTGGGT 2729
 QY 142 ----- 142
 Db 2730 CAATCTGGTCTGCCACTAACCTGACAGAGTGCCTCCCTGGGCTCAGTTTCTCTCT 2789
 QY 142 ----- 142
 Db 2790 CTGGAACAGGGATTAAGCAATAGCCCTGAGGTAGTGGTCAAGATTCAGTGGGACT 2849
 QY 142 ----- 142
 Db 2850 TGGATAGACAGCCCTGGCATGGCTTGTAGAGACTGGATCCTGACAGCTTGGTGACCCA 2909
 QY 142 ----- 142
 Db 2910 TGAAGTCCCGCCAGCCATGGCTTGGACAATATCTCCCAAGTGAAGTTGAAGAGCGCC 2969
 QY 142 ----- 142
 Db 2970 ACTCAGCCCTCCCTGGGTTGCTCAFTTACCAGGTTTGTAAACCGCAGCAGCAGCTG 3029
 QY 142 ----- 142
 Db 3030 GACCAGAGGGTGGGAAGAGCTGGGCTTTTATGCTTTTGGTTTTTAACTGGGC 3089
 QY 142 ----- 142
 Db 3090 AGAACAACACTGTGTTAGGGGCTAAGCTATGGAATAACCTTAGATGATGGTGAGGG 3149
 QY 142 ----- 142
 Db 3150 CCTGFGGCTGTGGCCCGAGGAGGAGTGTGCGAGCCACTGTFAGTACCGAGGCAG 3209
 QY 142 ----- 142
 Db 3210 AGAGTTCGTCTTAGTCCAAAGGCCACCCTCAAAGCGTGGCCCTGGGCTAGGATCCT 3269
 QY 142 ----- 142

Db 3270 GGGTGGGCTCAGTGGAGCTGGTTAAACAAGGTGAATCTTTGCGCTGTCCAGACTACT 3329
 QY 142 ----- 142
 Db 3330 GCATCGGAATCTATGATTTATCCAGGGTGTGTGCACATTAACAAGAGGCTTGAAGA 3389
 QY 142 ----- 142
 Db 3390 TCTGTGCCACCGCTCTCGTCAATATGACAGCCCTCTCCGTTACCCCTTAATCCCTTT 3449
 QY 142 ----- 142
 Db 3450 AAACACTCCACAGATCTTTCAGTGGTGGGACCTGGGGATGGCTGTCTCACCTAGGC 3509
 QY 142 ----- 142
 Db 3510 AGGTTCCACTAAGGGGCTTCCAGGGCTGCCGTAACCCCTGCAAGGACTTTCCTCAGCTC 3569
 QY 142 ----- 142
 Db 3570 TGAATGGCTGAGACCTTTCATTTTAAAGAACTTGGAGCCTGAGTAAAGCCTCATCTTT 3629
 QY 142 ----- 142
 Db 3630 TTTTTTTTTTTTTTAAATTTTTTAAATAATTTGGAGTTTCGCTCTTGTGGCCAGGCT 3689
 QY 142 ----- 142
 Db 3690 GGAGTCAATGFGGATCTGGCTCACCGCAACCTCCACCTCCAGGGTTCAAGTGAATC 3749
 QY 142 ----- 142
 Db 3750 TCCTCCCTCAACTCTCTGATAGCTGGGATTAACAGATGTGCCACTAGCCAGCTAAT 3809
 QY 142 ----- 142
 Db 3810 TTTGTATTTTTAGTAGATGGGTTTTCTCCCTGTTGGTTCAGGCCAGTCTCAAACTCCA 3869
 QY 142 ----- 142
 Db 3870 ACCTCTGFGATCTGCCCGCTCCGCTCCAGAGTGTGGGATTCAGGCAATGAGCCAC 3929
 QY 143 ----- ValTyrIleSerGluIleAlaTy 150
 Db 3930 TGGGCCAGCCATAAATCCTCATCTGATTTGGTGGCAGGTCTACATCTCGAAATCGCCTA 3989
 QY 150 rProAlaValArgGlyLeuLeuGlySerCysValGlnLeuMetValValGlyIleLe 170
 Db 3990 CCCAGCAGTCCGGGGTTCCTGGCTCTGTGTGCAGCTAAATGGTCTGTGGCATCCT 4049
 QY 170 uLeuAlaTyrIleu----- 174
 Db 4050 CCTGGCTACCTGGCAGGATAGTTGTCAATTAATCTTTCCTGTTTGTGGTCAAGTG 4109
 QY 174 ----- 174
 Db 4110 GCCATTTACAGATGTAAACAGAGTTTCAGTGGGTCAGGGACTTGTCTGAGGTCAC 4169
 QY 174 ----- 174
 Db 4170 ATAGCCAGTCCCAACATCCAGGCTCCCTTCTCCGGGACCTTGGGATAGGTGAGTCTTT 4229
 QY 174 ----- 174
 Db 4230 TTTTTTTTTTTTTTGAGACAGATCCCGCTGTGTCCGCCAGGCTGGAGTGCAGTGGTGGC 4289
 QY 174 ----- 174
 Db 4290 TCTCTCAACCTCTGCCCTCCGGGTTCAAGCGATTTCTCTGTCTCAGCTTCCCAAGTAGC 4349
 QY 174 ----- 174

Db 4350 TGGGACTACAGCATGTGCCACCACACCTGGCTAAATGTTGTAATTTTTGGTAGAAAGG 4409
 Qy 174 ----- 174
 Db 4410 GGTTTACCCTGTTAGCCAGGATGCTTTGATCTCCTGACCTCATGAAACCCCAACCTC 4469
 Qy 174 ----- 174
 Db 4470 AGCTTCCCAAGTGTGGATTAACAGGTGTAGCCACCGCCCGCGATAGGTGTGC 4529
 Qy 174 ----- 174
 Db 4530 TTATGACTAGTTCCAGGGACAGCTGGAGAAGGTATACCCTCAGCTTTCTGAAAGCCCC 4589
 Qy 174 ----- 174
 Db 4590 CTCTAGGGGTGGGAAGGCTGACTTGTAGGTCAGTACCTTACTGTGCAATGCAAAAC 4649
 Qy 174 ----- 174
 Db 4650 CAGAGGCTCCAGAGGTGGTGGTCCCTGTAAAAAATCTTCAGCCCGGGCGATAAGGG 4709
 Qy 174 ----- 174
 Db 4710 CTGAGCAAGTCCCTGAGAAATCAGTGCAGATTCACCATGTCCGGCAGTTGCACGTG 4769
 Qy 174 ----- 174
 Db 4770 AAGGAGCCGGGTCTCTGGGACCCCGAGGGCAAAAGCTGGCCAGCATGAAGAACTTC 4829
 Qy 174 ----- 174
 Db 4830 AAGAGACAGGTGTCTCATTAAGAAAGCCAGGCTTTAGAGCCAGACCAGTGGCT 4889
 Qy 174 ----- 174
 Db 4890 CACACCTGTAATCCAGACTTTGGGAGCAGAGGTGGCGGATCACCTGAGGTCAGGAG 4949
 Qy 174 ----- 174
 Db 4950 TTGAGATCAGCTGGCCAAACATGTTGAAAACCCATTTCTATTTAAATFACAAAATAG 5009
 Qy 174 ----- 174
 Db 5010 CCAGGCGTGTGTGGCCACCTGTAATCCAGCTACTGGGGCTGAGCCACAGAAATT 5069
 Qy 174 ----- 174
 Db 5070 GCTTGAACCCAGAGGTGAGGTTTCAGTGAGCTGAGATCGGCCACTGCACCTCCAGTGT 5129
 Qy 174 ----- 174
 Db 5130 GGGCGACACAGTGAAGTCTGTCTCAAAAAGAAAAGAAAACAAGGCGAGGCTTCAATC 5189
 Qy 174 ----- 174
 Db 5190 TGACCATCTGAGTTGGGATAGCTCTCCTACGTGTGTGACCTTTGGCCAAAGTCAT 5249
 Qy 174 ----- 174
 Db 5250 ACCACCTTGAAAGCCTCAGTTTCTGTCTTGAATGGGCATAACCTCATCTAGTCCAG 5309
 Qy 174 ----- 174
 Db 5310 TACCTGCTAATAAATTAATGACATGAAAAGTGCCTAACAGTGGTTCCAAATGAGA 5369
 Qy 174 ----- 174
 Db 5370 ATTGTCAGGGATGACTTGAGAACCATGGGATCATCGGGGTAGAGCCCTCGGGGCCCT 5429
 Qy 174 ----- 174
 Db 5430 GACGGGCTATGGCCCTGGCCCAAGGGCACAAGCCCTCTGCATCTGTCCAGGGCGGATACAC 5489

Qy 174 ----- 174
 Db 5490 AGGCTTTGGTGTTCATTTTTTACCACCTAAACCACCTTAAACAAATTCAGAAATCTGCCA 5549
 Qy 174 ----- 174
 Db 5550 GCGATCTGGCTACTTTGGGAGACCTGAGTAAACCAAGCCCCCACCCTGTGTGCAGGTGTG 5609
 Qy 174 ----- 174
 Db 5610 ACATTCATCATTTCCGGGACAGGTACTGGGTGTCCACCACCCCGTGGCACTATCTTCAGC 5669
 Qy 174 ----- 174
 Db 5670 TCACTTAATATGTTGTTGTTCTCCTGCTTTCTGATCAATTTCTGTGAGTTTCG 5729
 Qy 174 ----- 174
 Db 5730 ATCCACTGGGTGGTGAACCTCTGTATAACTCCACAAAGTGACATGGTGTGGCACCTTTG 5789
 Qy 174 ----- 174
 Db 5790 CTTACATTTGGGCAATCATGTTTCAAACGACTAGACTTCATAGAACAAATGCTGACACC 5849
 Qy 174 ----- 174
 Db 5850 CCAGCACAGAAAATAAGGACATCATGGTACCTTAGAGGTATTTTTCCACTGGAA 5909
 Qy 174 ----- 174
 Db 5910 CCAGGATGGTAAGCCTTGGGCTTTAGGAAAAAATAGTAAATTAGAGCAGCAATGCCCCTAG 5969
 Qy 174 ----- 174
 Db 5970 ACTCATTTTCCAGGGCCCCCTTTGTAAGAGAACCTTTTAACTTATATAAACTAGGTA 6029
 Qy 174 ----- 174
 Db 6030 ATCTGCACCTTATGCTGTCTTTCATACGTCAGAGCAGCAAAAATCCTCTCACCTCTACT 6089
 Qy 174 ----- 174
 Db 6090 TTCATGTATACTTTGAATAGATAATGCAATCAAAATGACTCAACACTCAAGTCCAGAGTG 6149
 Qy 174 ----- 174
 Db 6150 GAGAGTGCAGTGAAGGCTTCTCCGCTCTGCTCCCTTTCTTGTCTGGGGAGTCCC 6209
 Qy 174 ----- 174
 Db 6210 CATTGCCCTTTCTTGGTAGCCTTCAGAGAGTCTGTTTCATATAGGGACATGCATATTTA 6269
 Qy 174 ----- 174
 Db 6270 TCCTCCCTCCCCACCTCTTTCACAAATGGTGCCACATCATACCATCGGGACACGACGA 6329
 Qy 174 ----- 174
 Db 6330 AGCACCTCATCTTCATTTTTTGTGTGTACATAAATGTTTTCGTTGAAACGGATGCAGCTAG 6389
 Qy 174 ----- 174
 Db 6390 TAAATGTCCCGTGTFCATACCACCTGTCCGCTTCGCGACTGGTGTGATTTGGCCCGTGC 6449
 Qy 174 ----- 174
 Db 6450 TGTGCTTCATGAACAGTTGTGATAATGCAATGTCTAGTGCAGTAAACATAGGGCAGATGT 6509
 Qy 174 ----- 174
 Db 6510 GCAGAGGCGAGGGCCCCCATCTCTGAAGCTGAGTCTTTGAGGCTGGCAAGGCCCGGCT 6569


```

; LENGTH: 1545
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-4731

Alignment Scores:
Pred. No.:      5,47e-54      Length:      1545
Score:          605.00      Matches:     156
Percent Similarity: 51.60%      Conservative: 86
Best Local Similarity: 33.26%      Mismatches:  187
Query Match:    24.62%      Indels:      40
DB:             4           Gaps:        9

US-09-886-954A-1 (1-477) x US-09-489-039A-4731 (1-1545)
QY 20 AlaProArgGlyArgArg-----ValPheLeuAlaAlaPheAlaAlaAla 34
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
154 TC GCCGGGACGCAACGACTGATACCCCGCGGATGAACCTGGTGTGTTCATTCACGCGCGGCG 213

QY 35 LeuGlyProLeuSerPheGlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeu 54
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
214 GTGGTGGCTTGTCTTTGGCTGGATATCGGGGTGATATCCGGAGGCTGGCCCTTTATA 273

QY 55 GlnArgAlaAlaProProAlaProArgLeuAspAlaAlaAlaAlaSerTirpPheGlyAla 74
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
274 ACCGACCATTTTACCTTATCCAGCCAGCTTCAGGAG-----TGGGTGGTTAGC 321

QY 75 ValValThrLeuGlyAlaAlaAlaGlyGlyValLeuGlyGlyTirpLeuValAspArgAla 94
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
322 AGCATGATGTGGGGGGGGGATAGGCGCCCTGTTTAACGGCTGGCTGTCTTTCCGCCCTT 381

QY 95 GlyArgLysLeuSerLeuLeuLeuCysSerValProPheValAlaGlyPheAlaValIle 114
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
382 GCGGTAATACAGCTGATGGCGGGGCGTACTCTTTTGTGGCCGCTCTATCCGATCC 441

QY 115 ThrAlaAlaGlnAspValTirpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCys 134
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
442 GCTTTTCCGCCAGCGTGGAGGTCGTGGTGGCCGCGTGGTGGTGGGGGGGCGGACGC 501

QY 135 GlyValAlaSerLeuValAlaProValTyrIleSerGluIleAlaIleAlaValAlaVal 154
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
502 GGGATTGCTTATACCGGCCGCTGCTACTCTCCGAGATGGCCAGAGAGAGAGAGAGAG 561

QY 155 GlyLeuLeuGlySerCysValGlnLeuMetValValGlyIleLeuLeuAlaIleTyrLeu 174
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
562 GGGAAATGATGACTATGATATGATGATGATGATGATGATGATGATGATGATGATGAT 621

QY 175 AlaGlyTirpValLeuGluTirp-----ArgTirpLeuAlaValLeuGly-----CysVal 190
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
622 TCCGATACCCGCTTTAGCTACGCGGTAACTGGCGCGCCATGCTGGGCGGTCTGGCGCTG 681

QY 191 ProProSerLeuMetLeuLeuLeuMetCysPheMetProGluThrProArgPheLeuLeu 210
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
682 CCGGCGGTGATCCCTATGCTGTGCTCTTTTGGCGCGGCGGTTCCTCGGGCATCTGCTGC 741

QY 211 ThrGlnHisArgArgGlnGluAlaMetAlaAlaIleuArgPheLeuTirpGlySerGluGln 230
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
742 GAGAGGAGGAGCCATATATCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 801

QY 231 GlyTirpGluAspProProIleGlyAlaGlnGlnSerPheHisLeu----- 245
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
802 AAGCGCGCGGAGAGCTTAAAGAGATCCGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 861

QY 246 AlaLeuLeuArg---GlnProGlyIleTyrLysProPheIleIleGlyValSerLeuMet 264
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
862 GGGTGTGTAAAGATCAATCGTAACGCGCGCGGCGGTTCCTCGGGCATCTGCTGCAG 921

QY 265 AlaPheGlnGlnLeuSerGlyValAsmAlaValMetPheTyrAlaGluThrIlePheGlu 284
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
922 GCGATGCAGCAATTCACCGGGCATCAATCATGACTACTATGCGCCCGCGCTATCTTAA 981

QY 285 GluAlaLysPhe-----LysAspSerSerLeuAlaSerValValGlyValIle 301
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db 8729 GTCAGGCTTTCGGGCCAAGGTGCCAGAACCAAGCTGGCCCTTCTCTGGAGCCGGG 8788
QY 325 ----- 325

Db 8789 CCTGCCCTCTCAGAGGCATCCAGCCATGCTGTGGCCCTGGCGGTTCCCGGGGTGC 8848
QY 326 -----GlyValValMetValPheSerThrSerAlaPhe 336
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8849 CCGGCCACTGGCGCCCTCCCGCTCGCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8908

QY 336 eGlyAlaTyrPheLysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSe 356
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8909 CGCGCCCTACTTCAAGCTGACCCAGGTTGGCCCTGGCAACTCTTCGACAGTGGCACTC 8968

QY 356 rAlaProValSerAlaGlnProValAspAlaSerValGlyLeuAlaIleTirpLeuAlaValGI 376
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8969 GCGGCTGTCTGTACAGACCTGTGTGATGATGATGATGATGATGATGATGATGATGAT 9028

QY 376 ySerMetCysLeuPheIleAla----- 383
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
9029 CAGCATGTCCCTTTTCAATGCG-CCGTAAGGGGGCCCTGTGGGAGGCTGGGCGAGGATGG 9087

QY 383 ----- 383
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
9088 AGGTGATCCCCCGGCCCTGTGTGGCGCTCCCGCAGGTGGGTCTTCCCAGCCCTCA 9147

QY 383 ----- 383
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
9148 CCTCTCCTCTCGTAGTTCCAAGCTGTGATTTCTAGCTCTGACCAATGAGCTGTGGGA 9207

QY 383 ----- 383
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
9208 ACTPAGCCCTCTGCGTTCCTTTCACTTAATCCCTTGAGCAGCCCTCGAGCAGGCCCTG 9267

QY 383 ----- 383
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
9268 TGTGACCCCATCTCAGACAGACAGCCCTGTCGTTGGAGGAGAGCTGGCGAGGCT 9327

QY 383 ----- 383
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
9328 TGGCAGAAAGCATGGAGCTGTGGGACCACTGGAATCCGGGACCCACAGCCCCACA 9387

QY 383 ----- 383
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
9388 GGTGGGAGCTGGTCTTGTGCTCTGCGTCCACAGAGAGAGAGAGAGAGAGAGAGAGAG 9447

QY 384 -----GlyPheAlaValGlyTirpGlyProIlePr 393
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
9448 TGATGCTGTCCTTGGCTCTGCTCCACAGGCTTTGGGTTGGGCTGGGGGCCATCCC 9507

QY 393 oTirpLeuLeuMetSerGluIlePheProLeuHisValLysGlyValAlaIleThrGlyIleCy 413
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
9508 CTGGCTCCTCATGTCAGAGATCTTCCCTCTGCAAGTCAAGGGGTGGCCAGAGGATCTG 9567

QY 413 sValLeuThrAsnTirpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMet 432
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
9568 CGTCCTCACCACACTGGCTCATGCTTCTCGTACACAGAGGTTTCAGCAGCCCTCATG 9625

RESULT 3
US-09-489-039A-4731
; Sequence 4731, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIORITY FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIORITY FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4731

```



```

Db 982 ATGGCGGGCTTACCACTACTGAACAGCAGATGATCGCCACCCCTGGTGGCGCTGACC 1041
Qy 302 GlnValLeuPheThrAlaValAlaLeuLeuLeuMetAspArgAlaGlyArgLeuLeu 321
Db 1042 TTATGTTGGCCACCTTTATGGGTTCACAGTGTAAAGCGGCGCAAGCCAGCG 1101
Qy 322 LeuValLeuSerGlyValMetValPheSerThrSerAlaPheGlyAlaTyrPheLys 341
Db 1102 CTGAAATCGCTTATAGCTGATGGCGCTGGGCATCTCTGGTGGGC-----TACTGC 1155
Qy 342 LeuThrGlnGlyProGlyAsnSerSerHisValAlaLeuSerAlaProValSerAla 361
Db 1156 CTGATCAGTTCCAAATGGCACC----- 1179
Qy 362 GlnProValAspAlaSerValGlyLeuAlaTyrLeuAlaValAlaGlySerMetCysLeuPhe 381
Db 1180 -----GCATCCAGGGCCTCTCTGGCTCCGTCGGCATGACCATGATGTT 1227
Qy 382 IleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePhe 401
Db 1228 ATTGGCGGTATCGATGAGCGCGCGGGTGGTGGTGGTCTCTGTCCGAGATCCAG 1287
Qy 402 ProLeuHisValysGlyAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAla 421
Db 1288 CCGCTAAATGCGCGACTTCGGTATCACCTGCTCGACCACCACCACTGGGTGCGAAC 1347
Qy 422 PheLeuValThrLysGluPheSerSerLeuMetGluValLeuValLeuArgProTyrGlyAlaPhe 441
Db 1348 ATGATCATCGGGCCACTTCTGAGCTGTGAGCGATGGCGCGCCCGCCACCTTC 1407
Qy 442 TrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGlu 461
Db 1408 TGCTCTACAGCGCTCAACCTGGCTTTATCGGCATCACCTTCTGGCTGATCCGGAA 1467
Qy 462 ThrLysGlyLysThrLeuGluIle 470
Db 1468 ACCAAGAATGCACCCCTCGACCAATT 1494

```

```

RESULT 4
US-09-489-039A-4762
; Sequence 4762, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ. ID NOS: 14342
; SEQ. ID NO. 4762
; LENGTH: 1431
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4762

```

```

Alignment Scores:
Pred. No.: 3,116-51 Length: 1431
Score: 578.50 Matches: 149
Percent Similarity: 48.74% Conservative: 84
Best Local Similarity: 31.17% Mismatches: 180
Query Match: 23.54% Indels: 65
DB: 4 Gaps: 10

```

```

US-09-886-954A-1 (1-477) x US-09-489-039A-4762 (1-1431)

```

```

Qy 27 PheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeuGlyTyr 46
Db 82 TTCCTGCTGTTCTCCCGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 141
Qy 47 SerSerProAlaIleProSerLeuGlnArgAlaAlaProAlaProAlaProAlaAsp 66

```

```

Db 142 ATTGGCGGTCTTACCCCTTTATGGCAATGAGTTC-----CAGATTTCCGCC 189
Qy 67 AlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyValAlaAlaAlaGlyValLeu 86
Db 190 CACACCCAGGAGTGGTGTGAGTCCATGATGTTGGGGCTGCCGTCGGCGGCGG 249
Qy 87 GlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeuLeuCysSerValPro 106
Db 250 ACGCGTGGCTCTTTCAAACTGGCGGCGAAAGAGCTGATGATCGGCGCCATCCCT 309
Qy 107 PheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMetLeuLeuGly 126
Db 310 TTGCTGGCGGTTCGCTGTTCTCGCGCGCGCAAACTCGAGATCCCTGTGGTTC 369
Qy 127 ArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaProValTyrIleSer 146
Db 370 CBTGCTGCTCGGGCTGGCGGTGGCTCATATACGGCTCCCGTGTATCTGTGTCG 429
Qy 147 GluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGlnLeuMetValVal 166
Db 430 GAAATCGCCCCAGAAAAATTCGGCGAGTATGATTTCCATGTACCAGTGTATGATCAC 489
Qy 167 ValGlyIleLeuLeuAlaTyrIleuAlaGlyTrpValLeuGlu-----TrpArg 182
Db 490 ATCGGGATCTTGGCGCTATCTCTGTACACCGCTTTCAGCTACAGCGCGCATGGCG 549
Qy 183 TrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuLeuMetCysPheMet 202
Db 550 TGGATGCTGGGGTATCATCATCCGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 609
Qy 203 ProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaAlaLeu 222
Db 610 CCGGACAGCCCGCTGGTTCGCCCGCAAACTCGCTTGTGATGCGGAAACCGCTGCTG 669
Qy 223 -----ArgPhe 224
Db 670 CTGCGCTGCGGATACCCAGCGCAAGCGAGCTCGATGMAATCCGTTGAAAGC 729
Qy 225 LeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGluGlnSerPheHis 244
Db 730 CTGAAGTAAACAGTCCGGCTGG----- 753
Qy 245 LeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIle---IleGlyValSerLeu 263
Db 754 ---TCGCTGTTAAAGACACAGCAACTCCCGCGGGTGTCTCCGCAATCCCTGCTG 810
Qy 264 MetAlaPheGlnGlnLeuSerGlyValAenAlaValMetPheTyrAlaGluThrIlePhe 283
Db 811 CAGGTGATGCAACAGTTCACCGGGATGAACGTCATCATGTACTACGCGCGGAAGATCTTT 870
Qy 284 GluGluAlaLysPheLysAspSerSer-----LeuAlaSerValValGlyVal 300
Db 871 GAGTGGCGGGTATGCCCAACACCACTGAGCAAAATGTGGGGACAGTGTGCTGGCTG 930
Qy 301 IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgLeu 320
Db 931 ACTAAGTGTGGCCACCTTTATGCCCATCGTCTGGTCCAGCGCTGGGCGCGTAACCG 990
Qy 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
Db 991 ACGCTGATCTTGGCTTATCTGTATG-----GCCGCGGGA----- 1026
Qy 341 LysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer 360
Db 1027 -----ATGGCGCTCTGGGTACCATGATGATGATGATGATGATGATGATGAT 1077
Qy 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu 380
Db 1078 -----CAGTACATCGCGCTCTGTGATGCTGTGATG 1107
Qy 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIle 400

```

Db 1108 TTCATGTCGGTTCGGCATGAGCGCGCCACTGATTTGGGTACTGTCTCCGAAATC 1167
 Qy 401 PheProLeuHisValLysGlyValAlaThrGlyLeCysValLeuThrAsnTrpLeuMet 420
 Db 1168 CAGCGCTGAAAGCGCGGCACTTCGGTATCACCTGCTCCACAGCAGCAACTGGATTGCC 1227
 Qy 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
 Db 1228 AACATGATTCGGCGCCACCTTCCTGACATGCTCAACTGCTGGCAGCGCCCAATACC 1287
 Qy 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
 Db 1288 TTCGCTGTACGGCGGCTGAACAGTGTCTGTTATCTCTGCTGACGCTGTGGCTGATCCCG 1347
 Qy 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHis---PheGluGlyValArg 477
 Db 1348 GAACCAAAACGCTCCGCTGGAAACATATTAACGTAACCTGATGACAGGTCGT 1401

RESULT 5
 US-09-023-655-1104
 ; Sequence 1104, Application US/09023655
 ; Patent No. 6607879
 ; GENERAL INFORMATION:
 ; APPLICANT: Cocks, Benjamin G.
 ; APPLICANT: Susan G. Stuart
 ; APPLICANT: Jeffrey J. Seilhamer
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
 ; TITLE OF INVENTION: EXPRESSION
 ; NUMBER OF SEQUENCES: 1508
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/023,655
 ; FILING DATE: HERewith
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Zeller, Karen J.
 ; REGISTRATION NUMBER: 37,071
 ; REFERENCE/DOCKET NUMBER: PA-0001 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 1104:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3915 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GENBANK
 ; CLONE: g183684
 ; US-09-023-655-1104

Alignment Scores: 1.1e-49 Length: 3915
 Pred. No.: 570.50 Matches: 152
 Score: 49.90% Conservative: 88
 Best Local Similarity: 31.60% Mismatches: 190
 Query Match: 23.22% Indels: 51

DB: 4 Gaps: 9
 US-09-886-954A-1 (1-477) x US-09-023-655-1104 (1-3915)
 Qy 26 ValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeuGly 45
 Db 270 CTGATATTGGCCATCACAGTTGCTACACTCGGCTCTTCCAAATTTGGCTACCAACACTGGG 329
 Qy 46 TyrSerSerProAlaIleProSerLeuGlnArg----- 56
 Db 330 GTCATCAATGCTCCTGAGAAAGATCAATAAGGAATTTATCAATAAAACCTTTGACGACCAAG 389
 Qy 57 ---AlaAlaProAlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaVal 75
 Db 390 GGAATATGCCCCACCTCTCAGGTGCTGCACAGTCTCTCTGCTCTCTGCTGTGGCCATA 449
 Qy 76 ValThrLeuGlyAlaAlaAlaGlyValLeuGlyTrpLeuValAspArgAlaGly 95
 Db 450 TTTTCGTCGGGGTATGATCGGCTCTTTCCGTCGGACTCTTCGTCACCGCTTTGGC 509
 Qy 96 ArgLysLeuSerLeuLeuLeuCysSerValProPheValAlaGly-----PheAla 112
 Db 510 AGCGCAATTCATGCTGATTTGCAACCTGTTGGCTGTCACCTGGTGGCTTTATGGGA 569
 Qy 113 ValIleThrAlaAlaAlaGlnAspValTrpMetLeuGlyGlyArgLeuLeuThrGlyLeu 132
 Db 570 CTGTGTAAGTACTAAGTCGGTTGAAATGCTGATCTCTGGTGGCTTTGTTATTGGCCCTC 629
 Qy 133 AlaCysGlyValAlaSerLeuValAlaProValTyrIleSerGluIleAlaTyrProAla 152
 Db 630 TTCTGCGGACTCTGCACAGGTTTGTGCCCATGATGATGGAGAGATCTGCCCTACTGCC 689
 Qy 153 ValArgGlyLeuLeuGlySerCysValGlnLeuMetValValValGlyIleLeuLeuAla 172
 Db 690 CTGCGGGTCCCTTTGGCCTCTCAACACAGCTGGCATCGTTCTTGGAAATCTGTGGCC 749
 Qy 173 TyrLeuAlaGly-----TrpValLeu-----GluTrpArgTrpLeuAlaValLeuGly 188
 Db 750 CAGATCTTTGGTCTGGAAATTCATCTTTGGTCTGAAAGACTATGGCGGCTGCTACTGGGT 809
 Qy 189 -----CysValProProSerLeuLeuLeuLeuMetCysPheMetProGluThrPro 206
 Db 810 TTTTACCATCTCTCTGCTATCTTACAAAGTGCAGCCCTTCCATTTTCCCTGAAAGTCCC 869
 Qy 207 ArgPheLeuLeuThrGlnHisArgArgGlnGlu---AlaMetAlaAlaLeuArgPheLeu 225
 Db 870 AGATTTTGTCTCATTAACAGAAAGAGAGAGAAATGCTAAGCAGATCTCCACGCGGTG 929
 Qy 226 TrpGlySerGluGlnGlyTrpGluAspProIleGlyAlaGluGlnSerPheHis--- 244
 Db 930 TGGGCGACCCAGGATGATCCCAAGACATCCAGAGATGAAAGATGAGATGACAGGATG 989
 Qy 245 -----LeuAlaLeuLeuArgGlnProGlyIleTyrLysPro 256
 Db 990 TCACAGAAAAGCAAGTCACCGTCTAGAGCTTTAGAGTGTCCAGCTACCGACAGGCC 1049
 Qy 257 PheIleIleGlyValSerLeuMetAlaPheGlnLeuSerGlyValAsnAlaValMet 276
 Db 1050 ATCATCAITTCATTTGCTCTCAGCTCTCTCAGCAGCTCTCTGGGATCAATGCTGTGTT 1109
 Qy 277 PheTyrAlaGluThrIlePheGluAlaLysPheLysAspSerSerLeuAlaSerVal 296
 Db 1110 TATTACTCAACAGGAATCTTCAAGATGACAGGTTTCAAGCCCATCTATGCCACCATC 1169
 Qy 297 ValValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAla 316
 Db 1170 GCGCGGGTGTGGTTAATAACTATCTTCACTGTAGTTTCTCTATTTCCTGGTGGAAAGGCA 1229
 Qy 317 GlyArgArgLeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPhe 336
 Db 1230 GGAAGAAAGGACTCTGCATATGATA----- 1253
 Qy 337 GlyAlaTyrPheLysLeuThrGlnGlyProGlyAsnSerHisValAlaIleSer 356

Db	1254	-----GGCCITGGAGGGATGCGCTTTTGGTCCACGCTC	1286	---
QY	357	AlaProValSerAlaGlnProValAspAlaSerValGlyLeuAlaLrpLeuAlaValGly	376	
Db	1287	ATGACTGTTTCTTTGTTATTAAGGATACTATAATGGATAGACTTGTCTGTAITGG	1346	
QY	377	SerMetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProIleTrpLeuLeu	396	
Db	1347	GGTATCTGCTCTTCTAGCTCTTTTGAATTTGGACCGCCCAATCCCTGGTTATT	1406	
QY	397	MetSerGluIlePheProLeuHisValysGlyValAlaThrGlyIleCysValLeuThr	416	
Db	1407	GTGGCCGAACCTTTCAGCCAGGGCCCGCCCGAGCTGCATGCGCCGCTGCTCC	1466	
QY	417	AsnTrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArg	436	
Db	1467	AACTGGACCTCCAACTTCTAGTCGATGCTCTCCCTCCGCTCCTCACTATTAGGA	1526	
QY	437	ProTyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeu	456	
Db	1527	GGCTAC--GTTTTTATTATCTTACCCTTCCCTCATTTACCTTCTTGGCTTTTACCTTC	1583	
QY	457	PheCysValProGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGly	476	
Db	1584	TTCAAGTCCCTGAGACCGCTGCGAGGACTTTTGAGGATATCACAGGGCCTTTGAAGG	1643	
QY	477	Arg 477		
Db	1644	CAG 1646		

RESULT 6
US-09-679-686B-17
; Sequence 17, Application US/09679686B
; Patent No. 6624343
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Hightner, Jonathan E.
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: HEXOSE CARRIER PROTEINS
; FILE REFERENCE: B1160 US NA
; CURRENT APPLICATION NUMBER: US/09/679,686B
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: 60/081,131
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: PCT/US99/07561
; PRIOR FILING DATE: 1999-04-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-679-686B-17

Alignment Scores:
Pred. No.: 2,82e-47 Length: 1752
Score: 542.50 Matches: 150
Percent Similarity: 47.50% Conservative: 88
Best Local Similarity: 29.94% Mismatches: 202
Query Match: 22.08% Indels: 61
DB: 14 Gaps: 14

QY	17	GlySerAlaProArgGlyArg-----ValPheLeuAlaPheAlaAa	33	
Db	96	GGCGCAAGACTACCCTGGCAGGCTCACCCCTTCTTCTTCCACCTCGCTCGTGGCC	155	
QY	34	AlaLeuGlyProLeuSerPheGlyPheAlaLeuGlyTyrSerSerProAlaIle	51	
Db	156	GCCACCGGGCGCTCATCTTTGGATATGACATCGGTATCTCAGGGGGCGTTACGTC	215	

QY	52	ProSerLeuGlnArgAlaAaProProAla	61	
Db	216	RACCCTTCTCGAAAAGTCTTCCGGAGGTGTATGACAGAGCAGATGAAAGGCTCC	275	
QY	62	ProArgLeuAspAspAlaAaLalaSerTrpPheGlyAlaValValThr	77	
Db	276	GCCAGCCAGTACTGCAAGTACGACCAACAGCTGCTCCAGACCTTCACTTCCCTCTAC	335	
QY	78	LeuGlyAlaAaAlaGlyValLeuGlyTrpLeuValAspArgAlaGlyArgLys	97	
Db	336	CTCGGGCGCTCTCTCTTCTTTCGGCCGACCGTCCACCGTCTCTGGGCCCTTAAG	395	
QY	98	LeuSerLeuLeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAa	117	
Db	396	TGTCATCTTCCCGGGGCTCACCTTCTCATCGGGCTGCGCTTAAACGGGGCGGG	455	
QY	118	GlnAspValTrpMetLeuLeuGlyArgLeuLeuThrGlyLeuAlaCysGlyValAla	137	
Db	456	GAGAACATCGCCATCTCATCTCGACGCACTCTCTCGTCTCGGCTTGGCTTCGCC	515	
QY	138	SerLeuValAlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeu	157	
Db	516	AATCAGTCTGTCGGGTGACTCTGCGAGATGGCGCTCCGCGTCTCCGGGGCATGCTC	575	
QY	158	GlySerCysValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeu	174	
Db	576	RACATCGGGTTCAGCTCATGATCACCATCGGCATCTTGGCGGGCGGCTCATCAATTAC	635	
QY	175	AlaGlyTrpValLeuGluTrpArgTrpLeuAlaValLeuGly	188	
Db	636	GACACCAACAGATCAAGCCCGGTAC-----GGTGGCGCATCAGCCTGGCCATCGCG	689	
QY	189	CysValProProSerLeuMetLeuLeuMetCysPheMetProGluThrProArgPhe	208	
Db	690	GCCGTCCCGGGCGCATCATCACCCTGGGTCTTTTCTCCCGGACCCCAACTCC	749	
QY	209	LeuLeuThrGlnHisArgArgGlnGluAlaMetAlaLeuArgPheLeuTrpGlySer	228	
Db	750	CTCATCGAGGTGGCCACCCGGAGGGCGCGCCGCATCTCAACCGCATCCGGCGCAGC	809	
QY	229	Glu-----GlnGlyTrpGluAspProIleGlyAlaGlnSerPheHis	244	
Db	810	GACGTGGACATCAGCGAGGATACCGGACCTGCTGGTGGCGGAGGAGTCC-----	863	
QY	245	LeuAlaLeuLeuArgGlnPro-----GlyIleTyrIysPro---Phe	257	
Db	864	---AAGTGGTGCAGCACCCGFGCGCAACATCTCTGACGCGCAAGTACCGGGCCAGCTG	920	
QY	258	IleIleGlyValSerLeuMetAlaPheGlnLeuSerGlyValAsnAlaValMetPhe	277	
Db	921	ACCATGGCGATCATGATCCCTTCTTCAGCAGCTGACGGGGCATCAACGTCATCATGTC	980	
QY	278	TyrAlaGluThrIlePheGluAlaLysPheLys---AspSerSerLeuAlaSer---	295	
Db	981	TACGGCGGGTCTCTTCGAGACGCTGGGGTTCAAGGGCGACGCGCTGCTCATGTGGCC	1040	
QY	296	ValValValGlyValIleGlnValLeuPheThrAlaValAlaLeuIleMetAspArg	315	
Db	1041	GTCATACGGGGCTGGTCAACGTGTCGACGCTCGTGTTCCTGCTTCACTCGACCGG	1100	
QY	316	AlaGlyArgLeuLeuValLeuSerGlyValValMetValPheSerThrSerAla	335	
Db	1101	CTGGGTCCCGGAGACTGTTCTGCGGGCGGACGAGATGCTCTGAGCCAGCTGGGTG	1160	
QY	336	PheGlyAlaTyrPheLysLeuThrGlnGlyProGlyIleAsnSerHisValAlaIle	355	
Db	1161	GTGGCCACCTGATCGCGTCAAGTTCGGGACGAGCGGC-----	1199	
QY	356	SerAlaProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaVal	375	
Db	1200	-----GTGGGGGAGATGCC-----AAGGGGTACCGCGCGGCTGGGTG	1238	


```

Qy 285 GlnAlaLysPheLysAspSerSer-----LeuAlaSerValValValGlyValIleGln 302
Db 1104 ACGTGGGATTCAAGAAAGACCTTCTCATGTGCTCCCTCATCAGGGGCTCGTCAAC 1163
Qy 303 ValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArgLeuLeuLeu 322
Db 1164 GTCTTCCACCGCTGCTGTCATCGTCCAGCCGCTCGCCGCAAGCTGTTTC 1223
Qy 323 ValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPheLysLeu 342
Db 1224 CTCAGGGCGGGCGAGATGATGCTGTGCCAGCTCATCTGGGCGAGCTCATCGCGCC 1283
Qy 343 ThrGlnGlyProGlyAenSerSerHisValAlaIleSerAlaProValSerAlaGln 362
Db 1284 AAGTTCGGGACCGCGGACCGGG-----1307
Qy 363 ProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeuPheIle 382
Db 1308 -----GACATCCCAAGGGCTAGCGCGGCTCGTGTCTTTCATCTCGGCTAGCTC 1361
Qy 383 AlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePhePro 402
Db 1362 GCCGCTTCCCTCGTGGGGCCCTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1421
Qy 403 LeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPhe 422
Db 1422 CTGGAGATCCGCGCGGGCGGAGAGCATCAAGCTCCCGTCAACATGTTCTTCCCTTC 1481
Qy 423 LeuValThrLysGluPheSerSerLeuMetGluValLeuArgProFyrGlyAlaPheTrp 442
Db 1482 TGCATCCGCGAGGCTTCTCCATGCTCTGCCACTTCAAG---TTGGGCTCTTCTAC 1538
Qy 443 LeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluThr 462
Db 1539 TTCTTCCCGGCTGGGTCGTCATCAGCAGCGTCTTATTCGCCCTTCTTCTTCCGCGG 1598
Qy 463 LysGlyLysThrLeuGluGlnIle 470
Db 1599 AAGAACGTGCCATCCGAGGAGATC 1622

```

```

RESULT 8
US-09-679-686B-15
; Sequence 15, Application US/09679686B
; Patent No. 6624343
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: HEXOSE CARRIER PROTEINS
; FILE REFERENCE: B81160 US NA
; CURRENT APPLICATION NUMBER: US/09/679,686B
; CURRENT FILING DATE: 2003-01-16
; PRIORITY APPLICATION NUMBER: 60/081,131
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: PCT/US99/07561
; PRIOR FILING DATE: 1999-04-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1960
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (10)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (28)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:

```

```

; NAME/KEY: unsure
; LOCATION: (514)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (533)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (542)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (552)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (620)
; OTHER INFORMATION: n = a, c, g or t
; US-09-679-686B-15
Alignment Scores:
Pred. No.: 1.33e-44 Length: 1960
Score: 518.00 Matches: 148
Percent Similarity: 47.55% Conservative: 95
Best Local Similarity: 28.96% Mismatches: 198
Query Match: 21.08% Indels: 70
DB: 4 Gaps: 16
US-09-886-954A-1 (1-477) x US-09-679-686B-15 (1-1960)
Qy 7 GluGluThrGlnProLeuLeuGlyPro-----ProGly 17
Db 141 GAGAAAAAATCCCTCCGTTAGGAGGTATAGCAACGGAGGGGGCAGGAGTACCTTGG 200
Qy 18 GlySerAlaProArgGlyArgValPheLeuAlaAlaPheAlaAlaLeuLeuGlyPro 37
Db 201 AGCCTCACCCT-----TTTGTACCGTAAACATGTATAGTTGCAGCCATGGTGG 251
Qy 38 LeuSerPheGlyPheAlaLeuGlyTyrSerSerProAlaIle-----ProSerLeu 54
Db 252 TTAATCTTCGGTTACGATATCGGAATTTTCAGGTGGGTGCACATCCATGATCCGTTCTG 311
Qy 55 GlnArgAlaAlaProProAlaProArg-----63
Db 312 CTCAGTTTTTCCGTCGGTTCGGAAGAAGATTCGCAAAAACGGTGAACACAGTAC 371
Qy 64 -----LeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAla 81
Db 372 TGTCAAATACGACAGTCAGACACTGACGATGTTCCACGTCGCTGCTGTATCTCGCGCGTTG 431
Qy 82 AlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeu 101
Db 432 CTGTCTGCTGTTGTTGCTCCACCGTCACACGTAGGTTCCGCGGAAATTCCTCATGCTT 491
Qy 102 LeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrp 121
Db 492 TTCGGAGGCTTGTCTTCTCCGCGTCCCTTATCAACGGTTCCTTCCANCCANCCAGGTGG 551
Qy 122 MetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaLeuValAla 141
Db 552 NTGCTCATCGTGGTCCGATTTGCTCGGGTTCGGTATCCGGTTCCCAATCAGTCTGTG 611
Qy 142 ProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysVal 161
Db 612 CCACCTCTANCTATCTGAAATGGCTCCATACAAATATAGAGGACATTTGAACATGGCTTT 671
Qy 162 GlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu--- 180
Db 672 CAGTTGTCCATCAGTGTGGTATCTTGTGGCAATGTTGAACTTATTTCTTTGCTAAA 731
Qy 181 -----TrpArgTrp---LeuAlaValLeuGlyCysValProSerLeuMet 195

```

; APPLICANT: Li, Samuel X.
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: McNeill, Patricia D.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.455C11
 ; CURRENT APPLICATION NUMBER: US/09/643,597
 ; NUMBER OF SEQ ID NOS: 369
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 135
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-643-597-135

Alignment Scores: 5.71e-44 Length: 2856
 Pred. No.: 514.50 Matches: 145
 Score: 514.50 Conservative: 91
 Percent Similarity: 46.92% Mismatches: 168
 Best Local Similarity: 28.83% Indels: 99
 Query Match: 20.94% Gaps: 14
 DB: 3

US-09-886-954A-1 (1-477) x US-09-643-597-135 (1-2856)
 Qy 25 ArgValPheLeuAlaAalaPheAlaAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeu 44
 Db 210 CGCCTCACTGCTGTGGGAGGAGAGAGTGGCTTCCCTGCGGATTTGGGCTACAACACT 269
 Qy 45 Gly----- 45
 Db 270 GGAGTCATCAATGCCCGCCAGAGAGTGCAGAGGTTCTACACCAGACATGGGTCCAC 329
 Qy 46 ---TyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaProArgLeu 64
 Db 330 CGCTATGGGAGAGACTCCTGCCACC----- 356
 Qy 65 AspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaAlaGlyGly 84
 Db 357 ---AGCTCACCCAGCTCTGTGCTCCCTCAGTGGCCACTTTT-----TCGTGGGGGC 407
 Qy 85 ValLeuGly-----GlyTrpLeuValAspArgAlaGlyArgLeuSerLeu 100
 Db 408 ATGATGGCTCTCTCTGTGGGCTTTTCCTTAACCGCTTTGGCCGGGAGATTCAATG 467
 Qy 101 LeuLeuCysSer---ValProPheValAla-----GlyPheAlaValIleThr 115
 Db 468 CTGATGATGAACCTGTGCTCCCTGTGCTCCCGCTCATCGGGCTTCGAAACTG--- 524
 Qy 116 AlaAlaGlnAspValTrpMetLeuLeuGlyGlyArgLeuThrGlyLeuAlaCysGly 135
 Db 525 ---GCCAAGTCTTTGAGATGCTATCCTGGCCGCTTCATCCTCGGCTTACTGCGGC 581
 Qy 136 ValAlaSerLeuValAlaProValIleSerGluIleAlaTyrProAlaValArgGly 155
 Db 582 CTGACCACAGGCTTCCGCCATGATGTGGTGAAGTGCACCACAGCCTTTCTGTTGG 641
 Qy 156 LeuLeuGlySerCysValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAla 175
 Db 642 GCCCTGGGCGCCCTGCACAGCTGGGCATCGTCTGCGGCATCCTCATCGCCGAGGTTC 701
 Qy 176 GlyTrpValLeuGlu-----TyrArgTrpLeuAlaValLeu 187
 Db 702 GGC-----CTGGACTCCATATGAGCAAGAGACTGTGGCCCTGCTGTGAGCATC 755
 Qy 188 GlyCysValProSerLeuMetLeuLeuLeuMetCysPheMetProGluThrProArg 207
 Db 756 ATCTTCACTCCCGCCCTGCTGCAGTGATCGTCTGCTGCTGCTGCTGCTGCTGCTG 815
 Qy 208 PheLeuLeu---ThrGlnHisArgArgGlnGluAlaMetAlaAlaLeuArgPheLeuT 226

; APPLICANT: Wang, Aijun
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: McNeill, Patricia D.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.455C11
 ; CURRENT APPLICATION NUMBER: US/09/643,597
 ; NUMBER OF SEQ ID NOS: 369
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 135
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-643-597-135

Alignment Scores: 5.71e-44 Length: 2856
 Pred. No.: 514.50 Matches: 145
 Score: 514.50 Conservative: 91
 Percent Similarity: 46.92% Mismatches: 168
 Best Local Similarity: 28.83% Indels: 99
 Query Match: 20.94% Gaps: 14
 DB: 3

US-09-886-954A-1 (1-477) x US-09-643-597-135 (1-2856)
 Qy 732 ATCAAAGGTTGGGAGTGGAGTTGAGTTGGGAGGTGCTATGGTCCCTGCCCTTATA 791
 Qy 196 LeuLeuLeu---MetCysPheMetProGluThrProArgPheLeuLeuThrGlnHisArg 214
 Db 792 ATCCAGTAGGATCACTAGTCTCCAGACACTCCCAATCCATGATTGAAGGGGTTGAT 851
 Qy 215 ArgGlnGluAlaMetAlaAlaLeuArgPheLeuTrpGly-----SerGluGlnGly 231
 Db 852 CGCAGAGGCCAAGCTCAGCTTCAGAGAATTCGCGGCATCGCAATGTTGATGAAGAG 911
 Qy 232 TrpGluAspProIleGlyAlaGluGlnSerPheHisLeuAla----- 246
 Db 912 TTAATGACTTGTGGCAGCAAGTGAATTCCTAGCCAAAGTGGACACCCCTGGAGAAC 971
 Qy 247 LeuLeuArgGlnProGlyIleTyrIysProPheIle---IleGlyValSerLeuMetAla 265
 Db 972 TTTGTTCAAAAGAAG-----TACAGACCCCACTCACCATGGCAGTGTGATTCATTC 1025
 Qy 266 PheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGluThrIlePheGluGlu 285
 Db 1026 TTCCAGCACTCACTGGAATCAATGTCATGTTTACCGCCCGCTGTTGTCAGCTCC 1085
 Qy 286 AlaLysPheLysAspSerSer-----LeuAlaSerValValGlyValIleGlnVal 303
 Db 1086 ATCCGGTTTAGGATGACTCTCTAAATGTCAGCTGTGATCACCAGGGTTGTTAATGTT 1145
 Qy 304 LeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgLeuLeuLeuVal 323
 Db 1146 GTCCCAACTGTGCTCAATTTATGTTGTGCAAGTGGGTAGGAGAGCCCTTTTCCT 1205
 Qy 324 LeuSerGlyValValMetValPheSer-----ThrSerAlaPheGlyAlaTyr 339
 Db 1206 GAAGTGGAGTCAAAATGCTCAATTCGCCAGGCTGATGTCAGCTGCATGAGCAATGGAGCAAG 1265
 Qy 340 PheLysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProVal 359
 Db 1266 TTTGGA---ACTGATGGAAACCAGGT----- 1289
 Qy 360 SerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCys 379
 Db 1290 -----GATTTGCCAAAGTGTATGCAATTTGTTGGTCTCTTCATTTGC 1334
 Qy 380 LeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProIleProIleLeuMetSerGlu 399
 Db 1335 ATTATGATATCACTTTCCTGCTGCTAGTGGTCCCTAGCTGGTGTGCTAGTAGAG 1394
 Qy 400 IlePheProLeuHisValLysGlyValAlaIleThrGlyIleCysValLeuThrAsnTrpLeu 419
 Db 1395 ATCTTCCCTTGGAGATTCGTTCCAGCTGCTCAGAGTATCAATGTGTCCGTTGCAATGCT 1454
 Qy 420 MetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGly 439
 Db 1455 TTCACCTTCTTGTGATTCACCAAGTCTTTCGACAATGCTTTGCCACATGAAG---TTCCGC 1511
 Qy 440 AlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal 459
 Db 1512 TTGTTCTCTTCTTGTGCTTCTGTTGATCATGACATCTTCTCTACTTCTTCTTCT 1571
 Qy 460 ProGluThrLysGlyLysThrLeuGluGlnIle 470
 Db 1572 CCGGAAACAAAGGCGCATTCCAATTAAGAAATG 1604

RESULT 9
 US-09-643-597-135
 ; Sequence 135, Application US/09643597
 ; Patent No. 6426072
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Lijun
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Hosken, Nancy
 ; APPLICANT: Fanger, Gary R.

```

816 TTCTGCTCATCAACCGCAACGAGGAGAAACCGGGCCAGAGTGTGCTAAAGAAGTGGCC 875
227 GlySerGluGlnGlyTrpGluAspProPheGlyAlaGluGlnSerPheHis----- 244
876 GGGACAGTGTGAGTGTGACCCATCCCTGACGAGGAGTGAAGGAGAGAGTGGCAGATG 935
245 -----LeuAlaLeuLeuArgGlnProGlyIleTyrIleTyrIlePhe 257
936 CGGGAGAAGAGTCCATCCTGGAGCTGTTCCGGTCCCGCCCTACCGCCAGCCATC 995
258 lleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPhe 277
996 CTATCGTGTGTGTGTCAGCTGTCCACGAGCTGTCTGGCATCAACGCTGTCTCTAT 1055
278 TyrAlaGluThrIlePheGluAlaLysPheLysAspSerSerLeuAlaSerValVal 297
1056 TACTCCAGGACATCTCCAGAGGCGGGGGTGCAGACCCCTGTGTATCCACCATTGGC 1115
298 ValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGly 317
1116 TCCGGTATCGTCAACAGCGCTTCACTGTCGTCGCTGTTGTTGGTGGAGGAGAGGC 1175
318 ArgArgLeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGly 337
1176 CGGCGGACCCCTGCACCTCATA-----GGCCCTCGTGGCAGTGGCGGGTGTGCCATCTC 1196
338 AlaTyrPheLysLeuThrGlnGlyProGlyAenSerSerHisValAlaIleSerAla 357
1197 -----GGCCCTCGTGGCAGTGGCGGGTGTGCCATCTC 1232
358 ProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySer 377
1233 ACCATCGCCTAGCACTGTGGAGCAGCTACCTGGATGCTCTATCTGAGCATCTGGCC 1292
378 MetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMet 397
1293 ATCTTTGGCTTGTGGCTTCTTGAAGTGGTCTCTGGCCCATCCCATGGTTTCATCGTG 1352
398 SerGluIlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAen 417
1353 GCTGAACCTTTCAGCAGGGTCCAGCTCCAGTCCAGTGGCCATGGCTTGCAGGCTTCTCCA 1412
418 TrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgPro 437
1413 TGGACCTCAAAATTCATTTGGGATGCTTCCAGTATGTGGAGCACTGTGGTCCC 1472
438 TyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeu----- 453
1473 TAC-----GTCTTTCATCATCTTCCATGCTCTCTGTTCTGTTC 1511
454 -----PheThrLeuPheCysValProGluThrLysGlyLysThrLeuGluGlnIleThr 471
1512 TTCATCTCACCCTCAAAAGTTCCTGAGACTAAAGGCGGACCTTCCATGAGATCGCT 1571
472 AlaHisPhe 474
1572 TCCGGCTTC 1580

```

```

RESULT 10
US-09-480-884A-135
; Sequence 135, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalosh, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A

```

```

; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 135
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-480-884A-135

```

```

Alignment Scores:
Pred. No.: 5, 718-44 Length: 2856
Score: 514.50 Matches: 145
Percent Similarity: 46.92% Conservative: 91
Best Local Similarity: 28.83% Mismatches: 168
Query Match: 20.94% Indels: 99
DB: 4 Gaps: 14

```

```

US-09-886-954A-1 (1-477) x US-09-480-884A-135 (1-2856)
QY 25 ArgValPheLeuAlaAaPheAlaAaAlaLeuGlyProLeuSerPheGlyPheAlaLeu 44
Db 210 CGCCTCATGCTGGCTGTGGAGGAGCAGTGTGGCTCCCTGCAGTTGGCTACAACACT 269
QY 45 Gly----- 45
Db 270 GAGTGCATCAATGCCCCCGAAGGTGATCGAGGAGTTCACCAACGACATGGGTCCAC 329
QY 46 ---TyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaProArgLeu 64
Db 330 CGTATGGGAGAGCATCTCCGCCAC----- 356
QY 65 AspAspAlaAlaAaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaAlaGlyGly 84
Db 357 ---AGCTCACCCAGCTCTGGTCCCTCTCAGTGGCCATCTTT-----TCTGTTGGGGC 407
QY 85 ValLeuGly-----GlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
408 ATGATGGCTCCTCTCTGTGGCCCTTTTCGTTAAACCGCTTGGCGCGGAATCAATG 467
QY 101 LeuLeuCysSer---ValProPheValAla-----GlyPheAlaValIleThr 115
468 CTGATGATGAACCTGCTGGCTTCGTGTCGCCCGGCTCATGGGGTTCCTCGAAACTG 524
QY 116 AlaAlaGlnAspValTrpMetLeuLeuGlyArgLeuLeuThrGlyLeuAlaCysGly 135
525 ---GGCAAGTCTTGTGATGCTGATCTGGCCCGCTTCATCATCGTGTGTACTGGCC 581
QY 136 ValAlaSerLeuValAlaProValTyrIleSerGluIleAlaTyrProAlaValArgGly 155
582 CTGACCACAGGCTTGGTCCCATGTATGTGGTGAAGTGTCAACCCACAGCCCTTCTGGTGG 641
QY 156 LeuLeuGlySerCysValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAla 175
642 GCCCTGGGACCTGTCACAGCTGGGCATGCTGCTGGGCATCTCATCGCCCGAGGTTC 701
QY 176 GlyTrpValLeuGlu-----TrpArgTrpLeuAlaValLeu 187
702 GGC-----CTGGACTCCATCATGGCAACAGGACTGTGGCCCTGCTGTGAGCATC 755
QY 188 GlyCysValProProSerLeuMetLeuLeuMetCysPheMetProGluThrProArg 207
756 ATCTTCATCCGGCCCTGCTGCAGTGCATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 815
QY 208 PheLeuLeu---ThrGlnHisArgArgGlnGluAlaMetAlaAlaLeuArgPheLeuTrp 226
816 TTCCTGCTCATCAACCCGCAACGAGGAAACCGGGCCAAAGAGTGTCTAAAGAGTGGCC 875
QY 227 GlySerGluGlnGlyTrpGluAspProIleGlyAlaGluGlnSerPheHis----- 244
876 GGGACAGTGTGAGTGTGACCCATCCCTGAGGAGTGAAGGAGAGAGTGGCAGATGATG 935
QY 245 -----LeuAlaLeuLeuArgGlnProGlyIleTyrIleTyrIlePhe 257

```

936 CGGAGAAAGGTCACCATCTGGAGCTGTCGGCTCCCGGCTTCCCGCCACCCCATC 995
 258 IleIleGlyValSerLeuMetAlaPheGlnLeuSerGlyValAsnAlaValMetPhe 277
 996 CTCATCGCTGGTGGTCCAGCTGCCAGAGCTGTGGCATCAAGCGTCTTCTAT 1055
 278 TyrAlaGluThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValVal 297
 1056 TACTCCAGGACATCTTCGAGAAGCGGGGTGCAGCAGCCTGTGTATGCCACCATGGC 1115
 298 ValGlyValIleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGly 317
 1116 TCCGATATCTCAACAGCGCTTCTACTGTCGTGCTGTTGTGGTGGAGGAGCAGGC 1175
 318 ArgArgLeuLeuLeuLeuSerGlyValValMetValPheSerThrSerAlaPheGly 337
 1176 CGGGGAGCCCTGCACCTATA-----GGCCTCGCTGGCATGGCGGTTGTCCCATCTCANG 1232
 338 AlaTyrPheLysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAla 357
 1197 -----GGCCTCGCTGGCATGGCGGTTGTCCCATCTCANG 1232
 358 ProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTyrLeuAlaValGlySer 377
 1233 ACCATCGCGTACGACTGCTGGAGCAGTACCCTGGATGTCCTATCTGAGCATCTGGGCC 1292
 378 MetCysLeuPheIleAlaGlyPheAlaValGlyTyrGlyProIleProIleProIleLeuMet 397
 1293 ATCTTGGCTTGTGGCTTCTTGAAGTGGTCTGGCCCATCCCATGTTTCATCGTG 1352
 398 SerGluIlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsn 417
 1353 GCTGAATCTTTCAGCCAGGCTCCAGCTCCAGTCCCATGTCCTGGCTTCCCAAC 1412
 418 TyrLeuMetAlaPheLeuValThrLysGluPheSerSerSerLeuMetGluValLeuArgPro 437
 1413 TGGACCTCAAATTCATTTGGGCGATGCTTCCAGTATGTGGAGCAACTGTGTGTCC 1472
 438 TyrGlyAlaPheTyrLeuAlaSerAlaPheCysIlePheSerValLeu----- 453
 1473 TAC-----GTCCTCATCATCTTCACTGTGCTCTGTTCTGTTC 1511
 454 -----PheThrLeuPheCysValProGluThrLysGlyLysThrLeuGluGlnIleThr 471
 1512 TTCATCTTCACTTCAAGTTCCTGAGACTTAAGCCGGGACCTTCGATGAGATCGT 1571
 472 AlahisPhe 474
 1572 TCCGGCTTC 1580

RESULT 11
 US-09-542-615A-135
 ; Sequence 135, Application US/09542615A
 ; Patent No. 6518256
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Hosken, Nancy A.
 ; APPLICANT: Fanger, Gary R.
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
 ; FILE REFERENCE: 210121.455C8
 ; CURRENT APPLICATION NUMBER: US/09/542.615A
 ; CURRENT FILING DATE: 2000-04-14
 ; NUMBER OF SEQ ID NOS: 350
 ; SOFTWARE: Fast-Seq for Windows Version 3.0
 ; SEQ ID NO 135
 ; LENGTH: 2856
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-09-542-615A-135

Alignment Scores: 5.71e-44 Length: 2856
 Pred. No.: 514.50 Matches: 145
 Score: 46.92% Conservative: 91
 Percent Similarity: 28.83% Mismatches: 168
 Best Local Similarity: 20.94% Indels: 99
 Query Match: 4 Gaps: 14
 DB:

US-09-886-954A-1 (1-477) x US-09-542-615A-135 (1-2856)

25 ArgValPheLeuAlaAlaPheAlaAlaLeuLeuGlyProLeuSerPheGlyPheAlaLeu 44
 210 CGCTCATGCTGGCTGGGAGGAGCAGTGTGGTCCCTGCGCATGTTGGCTACAACT 269
 45 Gly-----
 270 GGATCATCAATGCCCCCAAGAGTATCAGAGGTTCTAACACAGACATGGTCCAC 329
 46 ---TyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProAlaProArgLeu 64
 330 CGCTATGGGAGAGCATCCGCCACC-----
 65 AspAspAlaAlaAlaSerTyrPheGlyAlaValValThrLeuGlyAlaAlaAlaGlyGly 84
 357 ---ACGCTCACCGCTCTGGTCCCTCAGTGGCCATCTTT-----TCTGTGGGGGC 407
 85 ValLeuGly-----GlyTyrLeuValAspArgAlaGlyArgLysLeuSerLeu 100
 408 ATGATGGCTTCTCTGGGGCTTTCGTTAACCGCTTGGCCGCGGAAATTCATG 467
 101 LeuLeuCysSer---ValProPheValAla-----GlyPheAlaValIleThr 115
 468 CTGATGATGAACTGCTGGCTTCTGGCCGCTGCTCATGGCTTCTCGAAACTG--- 524
 116 AlaAlaGlnAspValTyrMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGly 135
 525 ---GGCAAGTCTTGTAGATGCTCATCTGGGCGCTTCAFCATCGGTGTACTGGCGC 581
 136 ValAlaSerLeuValAlaProValTyrIleSerGluIleAlaTyrProAlaValArgGly 155
 582 CTGACCACAGCTTCGTCCTCATGTGTGGTGAAGTGCACCCACAGCCTTTCGTGGG 641
 156 LeuLeuGlySerCysValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAla 175
 642 GCCCTGGCACCTTCACAGCTGGCATCGTCTGGCATCTCATCGCCACCGAGTGTTC 701
 176 GlyTyrValLeuGlu-----TrpArgTyrLeuAlaValIle 187
 702 GGC-----CTGGACTCCATCATGGGCAACAGGACCTGTGGCCCTCTGCTGAGCATC 755
 188 GlyCysValProProSerLeuMetLeuLeuLeuMetCysPheMetProGluThrProArg 207
 756 ATCTTCATCCCGGCTGCTGAGTGCATCGTGTGGCTTCTGCCCCGAGAGTCCCGCC 815
 208 PheLeuLeu---ThrGlnHisArgArgGlnGluAlaMetAlaAlaLeuArgPheLeuTyr 226
 816 TTCTTGTCTCATCAACCGCAACGAGGAGAACCGGGCAAGAGTGTCTAAAGAGTGGCC 875
 227 GlySerGluGlnGlyTyrGluAspProIleGlyAlaGluGlnSerPheHis----- 244
 876 GGGACAGCTGAGCTGACCTGACCTGAGGAGATGAAGGAGAGAGTCCGAGATGATG 935
 245 -----LeuAlaLeuLeuArgGlnProGlyIleTyrLysProPhe 257
 936 CGGGAGAGAGAGGTCACCATCTCTGGAGTGTTCCTCCCGCTCCCGCCACTCCGCGAGCCATC 995
 258 IleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPhe 277
 996 CTCATCGCTGGTGGTGGAGGCTGTCAGCAGCTGTCGGCATCAAGCGTCTTCTAT 1055
 278 TyrAlaGluThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValVal 297


```

Db 1056 TACTCCACGAGCATCTTCGAGAAAGGGGGGGTGCAGCGCTGTGATGCCACCATTTGGC 1115
QY 298 ValGlyValIleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGly 317
Db 1116 TCGGGTATCGTCAACAGCGCTTCACTGTCGTGCTGCTGCTGGTGGGAGGAGCGC 1175
QY 318 ArgArgLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGly 337
Db 1176 CGCGGACCCCTGCACCTCATA----- 1196
QY 338 AlaTyrPheIleLeuThrGlnGlyProGlyAmsSerSerHisValAlaIleSerAla 357
Db 1197 -----GGCTCGTGGCATGGCGGGTTGTGCCATPACTCATG 1232
QY 358 ProValSerAlaGlnProValAspAlaSerValGlyLeuAlaIleValGlySer 377
Db 1233 ACCATCGCGCTAGCAGCTGCGAGCAGTACCCTGGATGCTCCTATCTGACATCGTGGCC 1292
QY 378 MetCysLeuPheIleAlaGlyPheAlaValGlyTyrGlyProIleProIleTrpLeuLeuMet 397
Db 1293 ATCTTTGGCTTTGGCGCTTCTTTGAAGTGGTCTGCGCCCATCCCATGGTTTCATCGTG 1352
QY 398 SerGluIlePheProeuhisValysGlyValAlaThrGlyIleCysValLeuThrAsn 417
Db 1353 GGTAACTCTTTCAGCCAGGGTCCACGTCCAGTGCATGCGGTTGCGAGGCTTCTCCAAC 1412
QY 418 TrpLeuMetAlaPheLeuValThrIysGluPheSerSerLeuMetGluValLeuArgPro 437
Db 1413 TGGACCTCAAAATTCAATTGGGCATGCTCCAGTATGTGGAGCACTGTGTGGTCC 1472
QY 438 TyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeu----- 453
Db 1473 TAC-----GTCTTCATCATCTTCACTGTGCTCTGCTGTTCTGGTTCNGTTC 1511
QY 454 -----PheThrLeuPheCysValProGluThrIysGlyIleValThrLeuGluIleThr 471
Db 1512 TTCATCTTCACCTACTTCAAAGTCTGAGACTTAAAGGCGGAGCCCTTCGATGAGATCGCT 1571
QY 472 AlaHisPhe 474
Db 1572 TCGGGCTTC 1580

```

```

RESULT 12
US-09-606-421B-135
; Sequence 135, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hoeken, Nancy R.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 135
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-606-421B-135
Alignment Scores:
Pred. No.: 5,71e-44 Length: 2856
Score: 514.50 Matches: 145
Percent Similarity: 46.92% Conservative: 91

```

```

Best Local Similarity: 28.83% Mismatches: 168
Query Match: 20.94% Indels: 99
DB: 4 Gaps: 14
US-09-886-954A-1 (1-477) x US-09-606-421B-135 (1-2856)
QY 25 ArgValPheLeuAlaAlaPheAlaAlaLeuLeuGlyProLeuSerPheGlyPheAlaLeu 44
Db 210 CGCCTCATGCTGGCTGGGAGGAGCAGTGGCTGGCTCCCTGCGAGTTTGGCTACACACT 269
QY 45 Gly----- 45
Db 270 GGAGTCAATCAATGCCCCCGAAGGTGATCGAGGAGTTCTACAACCCAGCATGGTCCAC 329
QY 46 ---TyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaProArgLeu 64
Db 330 CCCTATGGGAGAGCATCTGCCACC----- 356
QY 65 AspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaAlaGlyGly 84
Db 357 ---ACGCTCACCCAGCTGTCCTCCCTCCTCAGTGGCATCTTT-----TCGTGGGGGC 407
QY 85 ValLeuGly-----GlyTrpLeuValAspArgAlaGlyArgIysLeuSerLeu 100
Db 408 ATGATGGCTCTCTCTGTGGGCGCTTTTCGTAAACCGCTTTGGCGGGGAATTCAAATG 467
QY 101 LeuLeuCysSer---ValProPheValAla-----GlyPheAlaValIleThr 115
Db 468 CTGATGATGAACCTGTCGGCTTCTGTCGGCGTGTCCATGGGCTTCTGAAACTG--- 524
QY 116 AlaAlaGlnAspValTrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGly 135
Db 525 ---GCCAAGTCTTTGAGATGCTGATCTCTGGCCCGCTTTCATCATCGTGTGTACTGGCG 591
QY 136 ValAlaSerLeuValAlaProValTyrIleSerGluIleAlaTyrProAlaValArgGly 155
Db 582 CTGACCACAGGCTTCGTGCCCATGTATGTGGTGAAGTGTCAACCACACAGCTTTCGTTGG 641
QY 156 LeuLeuGlySerCysValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAla 175
Db 642 GCCCTGGGCACCCCTCACAGCTGGCCTGTCGTGGCAGTCTTCATCCGCCAGGTGTTTC 701
QY 176 GlyTrpValLeuGlu-----TrpArgTrpLeuAlaValLeu 187
Db 702 GGC-----CTGGACTCCATCATGGCAACAGGAGCTGTGGCCCTCTGTGTGAGCATC 755
QY 188 GlyCysValProProSerLeuMetLeuLeuLeuMetCysPheMetProGluThrProArg 207
Db 756 ATCTTTCATCCCGGCCCTGTCGTCAGTGCATCGTGTGCCCTTCTGCCCGAGAGTCCCGC 815
QY 208 PheLeuLeu---ThrGlnHisArgArgGlnGluAlaMetAlaAlaLeuArgPheLeuTrp 226
Db 816 TTCTGCTCATCAACCGCAACGAGAGAACCGGGCCAAAGGTGCTAAANGAAGTGTCCGC 875
QY 227 GlySerGluGlnGlyTrpGluAspProIleGlyAlaGluGlnSerPheHis----- 244
Db 876 GGGACAGCTGACGTGCCATCCATGACCTCGAGAGATGAAGGAGAGTGGCAGATGATG 935
QY 245 -----LeuAlaLeuLeuArgGlnProGlyIleTyrIysProPhe 257
Db 936 CGGGAGAAGAAGGTCCATCATCTGGAGTGTTCCTCCCGCCCTTACCAGCCAGCCCATC 995
QY 258 IleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPhe 277
Db 996 CTCATCGTGTGGTCTGAGCTGCCAGCAGTGTCCAGCAGTGTCTGGCATCAACGCTGTTCTAT 1055
QY 278 TyrAlaGluThrIlePheGluGluAlaIlyPheLysAspSerSerLeuAlaSerValVal 297
Db 1056 TACTCCACGAGCATCTTCGAGAAGCGGGGGTGCAGCAGCCTGTGTATGCCACCATTTGGC 1115
QY 298 ValGlyValIleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGly 317
Db 1116 TCCGGTATCGTCAACACACAGGCTTCACTGTGTGGTCTGCTGTTTGGTGGAGCGGAGCGC 1175

```

318 ArgArgLeuLeuValValLeuSerGlyValValMetValPheSerThrSerAlaPheGly 337
1176 CGCGCGCCCTGCACCTATA----- 1196
338 AlaTyrPheLysLeuThrGlnGlyProGlyProGlyAenSerSerHisValAlaLeuSerAla 357
1197 -----GCCCTCGCTGGCATGGGGTGTGCCATCTCATG 1232
358 ProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySer 377
1233 ACCATCGCGCTAGCACTGCTGGAGCAGCTACCCCTGGATGCTCTATCTGAGCATCGTGGCC 1292
378 MetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMet 397
1293 ATCTTTGGCTTTGTGGCTCTTTGAAGTGGTCTCTGGCCCATCCCAATGTTTCACTG 1352
398 SerGluIlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsn 417
1353 GCTGAACCTTCAGCCAGGCTCCAGTCCAGCTGCCATGGCGTTCAGGCTTCTCCAAC 1412
418 TrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgPro 437
1413 TGGACCTCAAATTCATTTGGGCATGCTTCCAGATATGGAGCAACTGTGTGTCC 1472
438 TyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeu----- 453
1473 TAC-----GTCCTTCATCATCTTCACGTGCTCTGGTCTGTTC 1511
454 -----PheThrLeuPheCysValProGluThrLysGlyLysThrLeuGluGlnIleThr 471
1512 TTCATCTTCACCTACTTCAAAGTTCCTGAGACTAAGGCGGACCTTCGATGAGATCGCT 1571
472 AlaHisPhe 474
1572 TCGGGTTC 1580

Qy 46 ---TyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaProArgLeu 64
330 CGCTATGGGAGAGCATCTGCCACC----- 356
65 AspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaAlaGlyGly 84
357 ---ACGCTCACCCAGCTCTGGTCCCTCTCAGTGGCCATCTTT-----TCTGTGGGGGC 407
85 ValLeuGly-----GlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
408 ATGATTGGCTCTCTCTGTGGGCTTTTCGTTAACCGCTTTCGCGCGGCGGAATCAATG 467
101 LeuLeuCysSer---ValProPheValAla-----GlyPheAlaValIleThr 115
468 CTGATGATGAACCTGCTGGCTTCGTCGCGCGTGTCTCATGGCTTCTCGAACTG--- 524
116 AlaAlaGlnAspValTrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGly 135
525 ---GGCAAGTCTTTGAGATGCTGATCTGGGCGCTTCATCFGGTGTGATCTGGCGC 581
136 ValAlaSerLeuValAlaProValTyrIleSerGluIleAlaTyrProAlaValArgGly 155
582 CTGACCACAGGCTTCGTCCTCATGTATGTGGTGAAGTGTCAACCACAGCTTTCGTTGG 641
156 LeuLeuGlySerCysValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAla 175
642 GCCCTGGCACCCTGCACCAGCTGGCATGCTGCTGGCATCTCTCATCGCCAGGTGTC 701
176 GlyTrpValLeuGlu-----TrpArgTrpLeuAlaValLeu 187
702 GGC-----CTGACTCCATCANGGCAACAGGACCTGTGGCCCTGTCTGAGCATC 755
188 GlyCysValProProSerLeuMetLeuLeuMetCysPheMetProGluThrProArg 207
756 ATCTTATCCCGCCCTGTGTCAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 815
208 PheLeuLeu---ThrGlnHisArgArgGlnGluAlaMetAlaAlaLeuArgPheLeuTrp 226
816 TTCCTGCTCATCAACCAACGAGGAAACCGGCGCAAGAGTGTCTAAAGAACTGCGC 875
227 GlySerGluGlnGlyTrpGluAspProIleGlyAlaGluGlnSerPheHis----- 244
876 GGGACAGTCACTGACCCATGACTGTCAGAGATGAAGAGAGAGTGTGCGCAGATGATG 935
245 -----LeuAlaLeuLeuArgGlnProGlyIleTyrLysProPhe 257
936 CGGGAGAAAGAGTCAACATCTGGAGCTGTCGCTCCCGCCGCTACCCGCGCAGCCATC 995
258 IleIleGlyValSerLeuMetAlaPheGlnLeuSerGlyValAsnAlaValMetPhe 277
996 CTCATCGCTGTGGTGTGCTGAGCTGTCCAGCAGCTGTCTGGCATCAACGCTGCTTCAT 1055
278 TyrAlaGluThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerVal 297
1056 TACTCCAGAGCACTTCAGAGAGCGGGGTGGAGCAGCTGTGTGTATGCCACCATGGC 1115
298 ValGlyValIleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGly 317
1116 TCCGGTATCGTCAACACCGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1175
318 ArgArgLeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGly 337
1176 CGGGGACCCCTGCACCTCATA----- 1196
338 AlaTyrPheLysLeuThrGlnGlyProGlyAenSerSerHisValAlaIleSerAla 357
1197 -----GCCCTCGCTGGCATGGGGTGTGCCATCTCATG 1232
358 ProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySer 377
1233 ACCATCGCGCTAGCACTGCTGGAGCAGCTACCCCTGGATGCTCTATCTGAGCATCGTGGCC 1292
378 MetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMet 397

RESULT 13
US-09-221-107-135
; Sequence 135, Application US/09221107
; Patent No. 6660838
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
; FILE REFERENCE: 210/21,455C2
; CURRENT FILING DATE: 1998-12-22
; NUMBER OF SEQ. ID NOS: 161
; SOFTWARE: PatentIn Ver. 2.0
; SEQ. ID NO 135
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Human
US-09-221-107-135
Alignment Scores:
Pred. No.: 5,71e-44 Length: 2856
Score: 514.50 Matches: 145
Percent Similarity: 46.92% Conservative: 91
Best Local Similarity: 28.83% Mismatches: 168
Query Match: 20.94% Indels: 99
DB: 4 Gaps: 14
US-09-886-954A-1 (1-477) x US-09-221-107-135 (1-2856)
Qy 25 ArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeu 44
Db 210 CGCCTCATGCTGGTGTGGAGGAGCAGTGTGGTCCCTGCGAGTTGGCTACACACT 269
Qy 45 Gly----- 45
Db 270 GGAGTCAATCAATGCCCCCAGAGGATGTCGAGGATTTTACAAACAGACATGGGTCCAC 329

```

Db 1293 ATCTTTGGCTTTGGCCCTTTTGAAGTGGTCTTGGCCCATCCATCGTTCATCGTG 1352
Qy 398 SerGluIlePheProLeuHisValLysGlyValAlaThrGlyLeuCysValLeuThrAsn 417
Db 1353 GCTGAACCTTTAGCCAGGTCACGCTCCAGCTGCAATGGCCGTTCCAGGCTTCTCCAAC 1412
Qy 418 TrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgPro 437
Db 1413 TGGACCTCAAATTTCAATTTGGGCATGTGCTTCCAGTATGTGGAGCAACTGTGTGTGCC 1472
Qy 438 TyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeu 453
Db 1473 TAC-----GTCCTTCATCATCTTCACTGTGTCTCCTGTTCTGTTTC 1511
Qy 454 -----PheThrLeuPheCysValProGluThrLysGlyThrLeuGluGlnIleThr 471
Db 1512 TTCATCTTACCTACTTCAAGTTCTGAGACTAAAGCCGGACCTTCGATGAGATCGCT 1571
Qy 472 AlaHisPhe 474
Db 1572 TCCGGCTTC 1580

```

RESULT 14

```

US-09-466-396A-135
; Sequence 135, Application US/09466396A
; Patent No. 6696247
GENERAL INFORMATION:
; APPLICANT: Hang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C4
; CURRENT APPLICATION NUMBER: US/09/466,396A
; CURRENT FILING DATE: 1999-12-17
; NUMBER OF SEQ ID NOS: 224
SOFTWARE: PastSeq for Windows Version 3.0
; SEQ ID NO 135
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-466-396A-135

```

Alignment Scores:

```

Pred. No.: 5,71e-44 Length: 2856
Score: 514.50 Matches: 145
Percent Similarity: 46.92% Conservative: 91
Best Local Similarity: 28.83% Mismatches: 168
Query Match: 20.94% Indels: 99
DB: 4 Gaps: 14

```

US-09-886-954A-1 (1-477) x US-09-466-396A-135 (1-2856)

```

Qy 25 ArgValPheLeuAlaPheAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeu 44
Db 210 CGCCTCATGCTGGCTGGAGGAGCAGTCTTGGCTCCCTGCGAGTTGGCTACAACT 269
Qy 45 Gly----- 45
Db 270 GAGTCAATGCCCCCAGAGGTATCGAGGATTTACACCAGACATGGGTCAC 329
Qy 46 ---TyrSerProAlaIleProSerLeuGlnArgAlaAlaProAlaProArgLeu 64
Db 330 CGCTATGGGAGAGCATCTGCCACC-----TCTGTTGGGGC 356
Qy 65 AspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaAlaGlyGly 84
Db 357 ---ACGCTCACCACGCTGCTGCTCCCTCCTCAGTGGCCATCTTT-----TCTGTTGGGGC 407
Qy 85 ValLeuGly-----GlyTrpLeuValValAspArgAlaGlyArgLysLeuSerLeu 100
Db 408 ATGATTTGGCTCTCTCTGTGGCCCTTTTCGTTAACCCGTTGGCCGCGGAATTCATG 467
Qy 101 LeuLeuCysSer---ValProPheValAla-----GlyPheAlaValIleThr 115

```

```

Db 468 CTGATGATGAACCTGTGGCCCTTCTGTCCGCGTCTCATGGCTTCTCGAAACTG--- 524
Qy 116 AlaAlaGlnAspValTrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGly 135
Db 525 ---GGCAAGCTCTTGGAGATGCTGTCCTGGGCGGCTTTCATCATCGGTGTACTCGGGC 581
Qy 136 ValAlaSerLeuValAlaProValTyrIleSerGluIleAlaLafyrProAlaValArgGly 155
Db 582 CTGACCACAGGCTCGTGGCCCATGTATGGGTGAAGTGTCAACACAGCCCTTTCGTGG 641
Qy 156 LeuLeuGlySerCysValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAla 175
Db 642 GCCCTGGGCACCTGACCAGCTGGGCATCGTCTGGCATCTCATCGCCACAGGTGTC 701
Qy 176 GlyTrpValLeuLuu-----TrpArgTrpLeuAlaValLeu 187
Db 702 GGC-----CTGACTCCATCATGAGGCAACAGGACCTGTGGCCCTGTCTGTGAGCAT 755
Qy 188 GlyCysValProSerLeuMetLeuLeuMetCysPheMetProGluThrProArg 207
Db 756 ATCTTCATCCCGCCCTGTGTGAGTGCATCGTGTCCCTTCTGCCCCGAGAGTCCCCGC 815
Qy 208 PheLeuLeu---ThrGlnHisArgArgGlnGluAlaMetAlaAlaLeuArgPheLeuTrp 226
Db 816 TTCTCTCATCAACCGCAACGAGGAAACCGGGCCAGAGTGTGCTAAAGAAGCTCGCG 875
Qy 227 GlySerGluGlnGlyTrpGluAspProIleGlyAlaGluGlnSerPheHis----- 244
Db 876 GGGACAGCTGACGTGACCATGACCTGCAGGAGATGAAGMAGAGAGTCCGCGAGATG 935
Qy 245 -----LeuAlaLeuLeuArgGlnProGlyIleTyrLysProPhe 257
Db 936 CGGGAGAGAGGTACCATCTGTGAGCTGTCGCTCCCGCTCCCGCTACCGCCAGCCATC 995
Qy 258 IleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPhe 277
Db 996 CTCACTCGCTGTGTGTGCGAGCTGCCAGCAGCTGTCTGGCATCAACGCTGTCTTCT 1055
Qy 278 TyrAlaGluThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerVal 297
Db 1056 TACTCCAGCATCTTCGAGAAAGCGGGGTGCACAGCCCTGTGTATCCACCATTGGC 1115
Qy 298 ValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGly 317
Db 1116 TCCGGTATCGTCAACACGGCTTCACTGTCTGTGCTGTGCTGTTGTGGTGGAGCGAG 1175
Qy 318 ArgArgLeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGly 337
Db 1176 CGGCGGACCCCTGCACCTCAT----- 1196
Qy 338 AlaTyrPheLysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAla 357
Db 1197 -----GGCCTCGCTGGCATGGCGGTTGTCCCATCTCATG 1232
Qy 358 ProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySer 377
Db 1233 ACCATCGCTAGCACTGCTGGAGCAGTACCCTGGATGCTCTATCTGAGCATCTGGCC 1292
Qy 378 MetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMet 397
Db 1293 ATCTTTGGCTTGTGGCTTCTTTGAAAGTGGTCTCTGGCCCATCCATCCATGGTTCATCG 1352
Qy 398 SerGluIlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsn 417
Db 1353 GCTGAACCTTTCAGCCAGGTCACAGTCCAGCTGCCATTCGCTTGGCCGTTCTCCAAC 1412
Qy 418 TrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgPro 437
Db 1413 TGGAACTCAATTTCAATTTGGGCGCATGCTTCCAGTATGTGGAGCAACTGTGTGTGCC 1472
Qy 438 TyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeu----- 453

```

Db 1473 TAC-----GTCATTCATCTTCCACTGTGCTCCTGCTGTCTGTTTC 1511
 Qy 454 -----PheThrLeuPheCysValProGluThrLysGlyLysThrLeuGluGlnIleThr 471
 Db 1512 TTCATCTTCCACTTCAAGTCTTGAGACTTAAGGCGGACCTTTCATGAGATCGCT 1571
 Qy 472 AlaHisPhe 474
 Db 1572 TCCGGCTTC 1580

RESULT 15
 US-09-476-496A-135
 ; Sequence 135, Application US/09476496A
 ; Patent No. 6706262
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Hosken, Nancy A.
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Fanger, Gary R.
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF
 ; FILE OF INVENTION: LUNG CANCER
 ; FILE REFERENCE: 210121.455CS
 ; CURRENT APPLICATION NUMBER: US/09/476,496A
 ; CURRENT FILING DATE: 1999-12-30
 ; NUMBER OF SEQ ID NOS: 254
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 135
 ; LENGTH: 2856
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-09-476-496A-135

Alignment Scores:
 Pred. No.: 5,71e-44 Length: 2856
 Score: 54.50 Matches: 145
 Percent Similarity: 46.92% Conservative: 91
 Best Local Similarity: 28.83% Mismatches: 168
 Query Match: 20.94% Indels: 99
 DB: 4 Gaps: 14

US-09-886-954A-1 (1-477) x US-09-476-496A-135 (1-2856)
 Qy 25 ArgValPheLeuAlaAalaPheAlaAalaLeuGlyProLeuSerPheGlyPheAlaLeu 44
 Db 210 CGCTCATCTGGCTGGAGGAGCAGTGTGGCTCCCTCCAGTTGGCTTACAACT 269
 Qy 45 Gly----- 45
 Db 270 GGAGTCAATCAATGCCCCAGAGGTGATCGAGGATTCACAACAGACATGGTCCAC 329
 Qy 46 ---TyrSerProAlaIleProSerLeuGlnArgAlaAlaProAlaProAlaProArgLeu 64
 Db 330 CGTATGGGAGAGCATCTGCCACC----- 356
 Qy 65 AspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaAlaGlyGly 84
 Db 357 ---ACGCTCACCAGCTGTGTCCTCCCTCAGTGGCATCTTT-----TCTGTGGGGC 407
 Qy 85 ValLeuGly-----GlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
 Db 408 ATGATGGTCTCTCTGTGGGCTTTTCGTTAACCGTTTGGCGGGGAATTCAAATG 467
 Qy 101 LeuLeuCysSer---ValProPheValAla-----GlyPheAlaValIleThr 115
 Db 468 CTGATGATCACTGGCTTGGTCCGGCTGCTCCGGCTCATGGCTTCTCGAACTG--- 524
 Qy 116 AlaAlaGlnAspValTrpMetLeuLeuGlyArgLeuLeuThrGlyLeuAlaCysGly 135
 Db 525 ---GGCAAGTCTTTCAGATGCTGATCTCCGGCGCTTTCATCATCGTGTGTACTCGGC 581
 Qy 136 ValAlaSerLeuValAlaProValTyrIleSerGluIleAlaTyrProAlaValArgGly 155
 Db 582 CTGACCACAGGCTTCGTGCCCATGATGTGGTGAAGTGTCAACCCACAGCCCTTCGTGGG 641

Qy 156 LeuLeuGlySerCysValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAla 175
 Db 642 GCCCTGGGACCCCTGCACCACTGGGCATCTCGTGGCATCTCCATCCATCCGCCAGGTGTT 701
 Qy 176 GlyTrpValLeuGlu-----TrpArgTrpLeuAlaValLeu 187
 Db 702 GGC-----CTGGACTCCATCATGGCAACCAAGGACCTGTGGCCCTCTCTGAGCATC 755
 Qy 188 GlyCysValProProSerLeuMetLeuLeuMetCysPheMetProGluThrProArg 207
 Db 756 AFTCTTATCCCGGCTCTGTCAGTGCATCGTGTCCCTCTGCCCCGAGAGTCCCCGC 815
 Qy 208 PheLeuLeu---ThrGlnHisArgArgGlnGlnAlaMetAlaAlaLeuArgPheLeuTrp 226
 Db 816 TTCTGTCTCATCAACCGCAACGAGGAGAACCGGCAAGAGTGTCTTAAGAGTGGC 875
 Qy 227 GlySerGluGlnGlyTrpGluAspProProIleGlyAlaGluGlnSerPheHis 244
 Db 876 GGGACAGCTGACGTGACCCATGACCTGCAGGAGATGAAGAGAGAGTGGCAGATGATG 935
 Qy 245 -----LeuAlaLeuLeuArgGlnProGlyIleTyrLysProPhe 257
 Db 936 CGGGAGAAGAAGTCACTCCTGGAGTGTTCGCTCCCGCCCTACCGCCAGGCCATC 995
 Qy 258 IleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPhe 277
 Db 996 CTCATCGTGTGGTCTGCTGACGCTGCCAGCAGCTGTCTGGCATCAACGCTCTCTAT 1055
 Qy 278 TyrAlaGluThrIlePheGluAlaLysPheLysAspSerSerLeuAlaSerValVal 297
 Db 1056 TACTCCAGCAGCATCTTCGAGAAGCGGGGTGCAGCAGCCTGTGTATGCCACCATGGC 1115
 Qy 298 ValGlyValIleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGly 317
 Db 1116 TCCGATATCGTCAACACCGCTTCACTGCTGCTGCTGTTGGTGGAGGAGCAGC 1175
 Qy 318 ArgArgLeuLeuLeuValSerGlyValValMetValPheSerThrSerAlaPheGly 337
 Db 1176 CGGGAGCCCTGCACCTCATA----- 1196
 Qy 338 AlaTyrPheLysLeuThrGlnGlyProGlyLysSerHisValAlaIleSerAla 357
 Db 1197 -----GGCCTCGCTGGCATGGGGTGTGCCATCATCATG 1232
 Qy 358 ProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySer 377
 Db 1233 ACCATCGGCTAGCATCTCTGGAGCAGTACCCTGGATGTCCTATCTGAGCATCTGGCC 1292
 Qy 378 MetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMet 397
 Db 1293 ATCTTTGGCTTTGGCTTCTTTGAAGTGGTCTGCGCCCATCCATGGTTCATCGTG 1352
 Qy 398 SerGluIlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsn 417
 Db 1353 GCTGAACCTTTCAGCAGGCTCCAGTCCAGTGCATTCGCTGGTTCAGGCTTCCCAAC 1412
 Qy 418 TrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgPro 437
 Db 1413 TGGACCTCAAAATTCATTTGGGCGCATGCTTCCAGTATGAGGCAACTGTGTGGTCCC 1472
 Qy 438 TyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeu----- 453
 Db 1473 TAC-----GTCCTTCATCATCTTCCACTGTCTCTCGTTCGTTC 1511
 Qy 454 -----PheThrLeuPheCysValProGluThrLysGlyLysThrLeuGluGlnIleThr 471
 Db 1512 TTCATCTTCACTTCAAAGTTCCTGAGACTTAAAGCCGAGCCTTCGATGAGATCGCT 1571
 Qy 472 AlaHisPhe 474
 Db 1572 TCCGGCTTC 1580

Search completed: February 25, 2005, 21:35:54
Job time : 276 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 25, 2005, 20:31:27 ; Search time 677 Seconds
(without alignments)
4174.955 Million cell updates/sec

Title: US-09-886-954A-1

Perfect score: 2457
Sequence: 1 MTPEDPEPTQLLPPGGSA.....CVPETKGTLEQITAHFEGR 477

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 5394803 seqs, 2962729879 residues

Total number of hits satisfying chosen parameters: 10789606

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DRV=xlp
-O=/cgn2_1/USPTO/spool_p/US09886954/runat_23022005_154244_13316/app_query.fasta_1.647
-DB=PublishedApplications_NA -QFM=fastap -SURFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blousum62
-TRANS=human40.cdi -LIST=45 -LOCAL=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09886954@cgn 1 1 723 @runat_23022005_154244_13316
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTV5_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Table with columns: Result No., Score, Query Match, Length, DB ID, Description. Contains 45 rows of sequence alignment data.

ALIGNMENTS

RESULT 1
US-10-115-831-137
Sequence 137, Application US/10115831
Publication No. US20030219743A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyang
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. US20030219743A1e1 Nucleic Acids and
FILE REFERENCE: 792CIPZADIV
CURRENT APPLICATION NUMBER: US/10/115,831
PRIOR FILING DATE: 2002-04-02
PRIOR APPLICATION NUMBER: 09/667,298
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 09/577,408
PRIOR FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 178

```

; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 137
; LENGTH: 1862
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4)..(1437)
US-10-115-831-137

Alignment Scores:
Pred. No.: 1,78e-241 Length: 1862
Score: 2457.00 Matches: 477
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-09-886-954A-1 (1-477) x US-10-115-831-137 (1-1862)
QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlySerAla 20
Db 4 ATGACGCCGAGACCCAGAGAAACCCAGCCGCTTCGGGGCTTCCTGGGGCAGCGCG 63
QY 21 ProArgGlyArgValPheLeuAlaPheAlaAlaLeuGlyProLeuSerPhe 40
Db 64 CCCCAGGGGGCGCGCTTCCTCGCCGCTTCGCGGCTTCGCGGCTTCGCGGCTTCAGCTTC 123
QY 41 GlyPheAlaLeuGlyTyrSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
Db 124 GGCTTCGGCTTCGGCTTCAGCTCCCGGCAATCCCTAGGCTCGAGCGCGCGCCCGG 183
QY 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
Db 184 GCCCGCGCGCTGGACGACGCGCGCGCTCCTCGTTCGGGGCTTCGCTGACCCCTGGGTC 243
QY 81 AlaAlaGlyValLeuGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
Db 244 GCAGCGGGGGAGTGTGGCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 303
QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
Db 304 TTCCTGCTCCGTCGCTCGGCGGCTTCGCGGCTTCGCGGCTTCGCGGCTTCGCGGCT 363
QY 121 TrpMetLeuGlyArgLeuLeuThrGlyValAlaCysGlyValAlaSerLeuVal 140
Db 364 TGGATGCTGGGGGGCGCCCTCCACCGGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCG 423
QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
Db 424 GCCCGGCTTACATCTCCGAAATCGCTACCCAGCAGTCCGGGGGTTGCTCGGCTCCTGT 483
QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrIleLeuAlaGlyTrpValLeuGlu 180
Db 484 GTGCAGCTAATGTGCTGCTGGCGGCTTCCTCGGCTTCCTCGGCTTCCTCGGCTTCG 543
QY 181 TrpArgTrpIleAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
Db 544 TGGCGCTGGCGGCTGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGG 603
QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
Db 604 TTCATGCCGAGACCCCGGCTTCCTGCTGACTCAGCACAGGGGCCAGGAGGCAATGGCC 663
QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProIleGlyAlaGlu 240
Db 664 GCCCTCGGCTTCCTGGGGCTCCGAGCGGCTGGGAGACCCCGGCTTCCTCGGCTTCG 723
QY 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIleTyrIleGly 260
Db 724 CAGAGCTTTCACCTGGCCCTCGTGGCGGCGCGGCTTCACAGCCCTTCATCATCGGT 783
QY 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280

```

```

784 GTCCTCCCTGATGGCCTTCAGCAGCTGTCCGGGGTCAAGCGGCTCATGTTCTATGCAGAG 843
QY 281 ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValValGlyVal 300
Db 844 ACCATCTTTGAAGAGGCCAAGTTCAAGGACAGCAGCCTGGCCTCGGTGCTGTGGGTGTC 903
QY 301 IleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArgLeu 320
Db 904 ATCCAGGTGCTGTTTACAGCTGTGGCGGCTTCATCATGCAACAGAGCAGGGCGGAGGCTG 963
QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
Db 964 CTCCTGGTCTTGTTCAGGTGTGGTTCATGTTTCAGCAGCAGGTCCTTCGGCCCTACTTC 1023
QY 341 LysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer 360
Db 1024 AAGTCAGCCAGGGTGGCCCTCGCAACTCCTCGCAGCTGGCCATCTCGCGGCTGTCTCT 1083
QY 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu 380
Db 1084 GCACAGCTGTTGATGCCAGCGTGGGCTGGCTGGCTGGCTGGCGGCGAGCATGTGCTTC 1143
QY 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProIleProIleLeuMetSerGluIle 400
Db 1144 TTCATCGCCGCTTTCGGTGGCTGGGCGCCATCCCTGGCTCCTCATGTGAGAGATC 1203
QY 401 PheProIleHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
Db 1204 TTCCTCTGATGTCAGGAGGGGCGGCGAGCAGGATCTGGCTCCTCACCACCTGGCTCATG 1263
QY 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
Db 1264 GCCTTTCGTGACCAAGAGGTTTCAGCAGCCTCATGGAGTCTCAGGGCTTCAGGGCC 1323
QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
Db 1324 TTCCTGGTTCCTCCGCTTCGATCTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1383
QY 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db 1384 GAACTAAGAAAGACTCTGGAACAATCACAGCCATTTTGGGGGGCGA 1434

RESULT 2
US-10-168-651-28
; Sequence 28, Application US/10168651
; Publication No. US20030171275A1
; GENERAL INFORMATION:
; APPLICANT: INCIYTE GENOMICS, INC.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: BURFORD, Neil
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: YANG, Junning
; APPLICANT: REDDY, Roopa
; APPLICANT: LAL, Preeti
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: YUE, Henry
; APPLICANT: NGUYEN, Damiel B.
; APPLICANT: YAO, Monique G.
; APPLICANT: TANG, Y. Tom
; APPLICANT: KHAN, Farrah A.
; TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS
; FILE REFERENCE: PI-0005 PCT
; CURRENT APPLICATION NUMBER: US/10/168,651
; PRIOR APPLICATION NUMBER: 60/172,000; 60/176,083; 60/177,332; 60/178,572; 60/179,758;
; 60/181,625
; PRIOR FILING DATE: 1999-12-23; 2000-01-14; 2000-01-21; 2000-01-28; 2000-02-02;
; 2000-02-10
; NUMBER OF SEQ ID NOS: 54

```



```

FEATURE:
NAME/KEY: misc_feature
LOCATION: (792)..(792)
OTHER INFORMATION: n equals a.t.g. or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1119)..(1119)
OTHER INFORMATION: n equals a.t.g. or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1156)..(1156)
OTHER INFORMATION: n equals a.t.g. or c
US-10-264-237-688
Alignment Scores:
Pred. No.: 4.52e-162 Length: 1156
Score: 1679.50 Matches: 344
Percent Similarity: 72.12% Conservative: 0
Best Local Similarity: 72.12% Mismatches: 10
Query Match: 68.36% Indels: 125
DB: 17 Gaps: 1
US-09-886-954A-1,(1-477) x US-10-264-237-688 (1-1156)
QY 1 MetThrProGlnAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
Db 53 ATGACGCCGAGACCAGAGAAACCCAGCCGCTTCGGGGCTCCTCGGGCCAR---- 108
QY 21 ProArgGlyArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
Db 108 ----- 108
QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
Db 108 ----- 108
QY 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValThrLeuGlyAla 80
Db 108 ----- 108
QY 81 AlaAlaGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgIleLeuSerLeu 100
Db 108 ----- 108
QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
Db 108 ----- 108
QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
Db 108 ----- 108
QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
Db 109 -----GTCTACATCTCCGAAATRCCTACCCAGCAGTCCGGGGTGTCTCGGCTCTGT 162
QY 161 ValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
Db 163 GTGCAGCTAATGTCGTCTGCGCATCTCTGCGCTACTGCGCCCTGCGAGCGTGGTGGAG 222
QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
Db 223 TGGCGTGGCTGGCTGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 282
QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnAlaMetAla 220
Db 283 TTCATGCCGAGACCCCGCTTCTCTGCTGACTCAGCAGCAGCGCCGAGGAGCCATGGCC 342
QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu 240
Db 343 GCCTCGGNTCTCTGTGGGGCTCCGAGCAGGGCTGGAGACCCCCCTCGGGCTGGAG 402
QY 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIlePheProPheIleIleGly 260

```

```

Db 403 CAGAGCTTTTCCCTGGCCCTGTGGCCANCCGGCATCTTACAAGCCCTTCATCATCGGC 462
QY 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
Db 463 GTCCTCCCTGATGGCTTCCAGCAGCTGTCCGGGGTCAAGCCGCTCATGTTCATGAGAG 522
QY 281 ThrIlePheGluAlaLysPheIleAspSerSerLeuAlaSerValValGlyVal 300
Db 523 ACCATCTTTGAAGAGGCCAAGTTCAAGGACAGCAGCCTGGCCCTCGCTCGTGGGTGC 582
QY 301 IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgIleArgLeu 320
Db 583 ATCCAGGTGCTTTCACAGCTGTGGCGCTCTCATCATGGACAGCAGCGGGGAGGCTG 642
QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
Db 643 CTCTCTGGTCTTCTCAGTGTGTGTCATGGTGTTCAGCAGCAGGTCCTTCGGGGCTTCT 702
QY 341 LysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer 360
Db 703 AAGCTGACCCAGGCTGGCCCTGGCAACTCTCGCAGCTGGCCATCTCGGGCCCTGTCT 762
QY 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValIleGlySerMetCysLeu 380
Db 763 GCACAGCCTGTTGATGCCAGCCTGGGGCTTNGCCTGGCTGGCGGAGCATGTGCTC 822
QY 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIle 400
Db 823 TTCATCGCCCGCTTTCGGGTGGGCTGGGGGGCCCAATCCCTGGCTCTCATGTGAGAGATC 882
QY 401 PheProLeuHisValIleValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
Db 883 TTCCTCTTTCATGTCAAGGGCTGGCGACAGCAWTTGCTCTCTCAMCAACTGGCTCATG 942
QY 421 AlaPheLeuValThrIleGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
Db 943 GSCCTTCTYGGKACCAAGAGTTCAGCAGCCTCATGGAGGTCCTCAGGGCCCTATGGAGCC 1002
QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
Db 1003 TTCCTGGCTTCCCTCGCTTTCAGTTCATCTTCAGTTCCTTTTCACTTTGGTCTGTC-CC 1061
QY 461 GluThrLysGlyThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db 1062 GAAACTAAAGAAAGAGACTCTGGAAACAATCACAGCCCAATTTTGGAGGGCGGA 1112

```

```

RESULT 5
US-10-169-395-29
Sequence 29, Application US/10169395
Publication No. US20040034192A1
GENERAL INFORMATION:
APPLICANT: KATO Seishi
APPLICANT: KIMURA, Tomoko
TITLE OF INVENTION: HUMAN PROTEINS HAVING HYDROPHOBIC DOMAINS AND DNAs ENCODING
FILE REFERENCE: 01997.015100.US
CURRENT APPLICATION NUMBER: US/10/169,395
CURRENT FILING DATE: 2002-11-29
PRIOR APPLICATION NUMBER: JP 2000-585
PRIOR FILING DATE: 2000-01-06
PRIOR APPLICATION NUMBER: JP 2000-588
PRIOR FILING DATE: 2000-01-06
PRIOR APPLICATION NUMBER: JP 2000-2299
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: JP 2000-26862
PRIOR FILING DATE: 2000-02-03
PRIOR APPLICATION NUMBER: JP 2000-58367
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: PCT/JP00/09359
PRIOR FILING DATE: 2000-12-28
NUMBER OF SEQ ID NOS: 150
SEQ ID NO 29

```

```

; LENGTH: 1461
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)..(849)
US-10-169-395-29

Alignment Scores:
Pred. No.: 46-149 Length: 1461
Score: 1554.50 Matches: 324
Percent Similarity: 67.8% Conservative: 1
Best Local Similarity: 67.64% Mismatches: 1
Query Match: 63.27% Indels: 153
DB: 17 Gaps: 2

US-09-886-954A-1 (1-477) x US-10-169-395-29 (1-1461)
Qy 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProGlyGlySerAla 20
Db 61 ATGACCCCGAGACCACAGGAAACCCAGCCGCTTCCTGGGGCCCTCCCTGGGGCGACGGCG 120
Qy 21 ProArgGlyArgValPheLeuAlaPheAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
Db 121 CCCCCGGCGCGCGCTTCCTCGCGCTTCGCGCTTCGCGCTTCGCGCTTCGCGCTTCGCGCTTC 180
Qy 41 GlyPheAlaLeuGlyTyrSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
Db 181 GGCTTCGCGCTCGCTACAGCTCCCGCGCATCCCTAGCTCGACGGCGCGCGCGCGCGCGCG 240
Qy 61 AlaArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
Db 241 GCCCGCGCGCTGACAGCCCGCGCTTCGCGCTTCGCGCTTCGCGCTTCGCGCTTCGCGCTTC 300
Qy 81 AlaAlaGlyValLeuGlyTyrLeuValAspArgAlaGlyArgIleLeuSerLeu 100
Db 301 GCGCGCGCGGAGTGTGTGGCGGCTGGCTGGTGGACCGCGCGCGCGCGCGCGCGCGCGCTC 360
Qy 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
Db 361 TTGCTGTGTCTCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTC 420
Qy 121 TrpMetLeuLeuGlyGlyArgLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
Db 421 TGGATGCTGTGGGGCGCGCGCTTCCTACCGCGCTGGCTGGGCTGGCTGGCTGGCTGGCTGG 480
Qy 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
Db 481 GCCCGGCTTACATCCGAAATCGCTACCCAGCAGTCCGGGGTGGCTGGCTGGCTGGCTGGCT 540
Qy 161 ValGlnLeuMetValValGlyIleLeuLeuAlaTyrIleLeuAlaGlyTyrValLeuGlu 180
Db 541 GTGCAGCTAATGTCGTCGCGCATCTCTGGCTTCCTGGCTTCCTGGCTTCCTGGCTTCCTGG 600
Qy 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
Db 601 TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 660
Qy 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAla-MetAl 220
Db 661 TTCATGCCGAGACCCCGCGCTTCCTGCTGACTCAGCACAGGCGCGCGCGCGCGCGCGCGCT 718
Qy 220 aAlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProIleGlyAlaGln 240
Db 718 ----- 718
Qy 240 uGlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIleGln 260
Db 718 ----- 718
Qy 260 yValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGln 280
Db 718 ----- 718

```

```

280 uThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVa 300
718 ----- 718
300 lIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLe 320
718 ----- 718
320 uLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPh 340
719 -----CTGGCTTCCTGACAGTGGTGTGTCAGTGGTTCAGCACAGAGTCCCTCGGCCCTTACTT 774
340 eLysLeuThrGlnGlyGlyProGlyAsnSerHisValAlaIleSerAlaProValSe 360
775 CRAAGCTGACCCAGGGTGGCCCTGGCAACTCTCCGACCTGGCCATCTCGGCGCGCTGTCTC 834
360 rAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLe 380
835 TGCACAGCTTGTGATGCCAGCGTGGCGCTGGCTGGCTGGCTGGCTGGCGACGATGTGCC 894
380 uPheIleAlaGly-PheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluI 400
895 CTTCACTCCCGGAGGTC----- 911
400 lPheProLeuHisValLysGlyValAlaAlaThrGlyIleCysValLeuThrAsnTrpLeuM 420
911 ----- 911
420 eAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuLeuArgProTyrGlyA 440
912 -----CTCAGGCCCTATGGAG 927
440 lAPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValP 460
928 CTTTCTGGCTTGGCTTCGCTTCGCACTTCAGTGTCCCTTTTCCATTTTCTCTGTGTCTCC 987
460 roGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
988 CTGAACATAAGGAAAGACTCTGGAAACAATCAGCCCAATCAGCCCAATTTTGGGGGGCGA 1040

```

```

RESULT 6
US-10-169-395-19
; Sequence 19, Application US/10169395
; Publication No. US20040034192A1
; GENERAL INFORMATION:
; APPLICANT: KATO, Seishi
; APPLICANT: KIMURA, Tomoko
; TITLE OF INVENTION: HUMAN PROTEINS HAVING HYDROPHOBIC DOMAINS AND DNAS ENCODING
; TITLE OF INVENTION: THESE PROTEINS
; FILE REFERENCE: 01997.015100.US
; CURRENT APPLICATION NUMBER: US/10/169,395
; CURRENT FILING DATE: 2002-11-29
; PRIOR APPLICATION NUMBER: JP 2000-585
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: JP 2000-588
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: JP 2000-2299
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-26862
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: JP 2000-58367
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: PCT/JP00/09359
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 150
; SEQ ID NO 19
; LENGTH: 789
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-169-395-19
Alignment Scores:

```


Db 1493 ACGGTCCTGGAGCAGATCGAGTCCTTCTCCGCACCGGGGAGA 1535

RESULT 8
 US-10-172-118-1685
 Sequence 1685, Application US/10172118
 Publication No. US20030224374A1
 GENERAL INFORMATION:
 APPLICANT: Dai, Hongyue
 APPLICANT: He, Yudong
 APPLICANT: Linsley, Peter
 APPLICANT: Mao, Mao
 APPLICANT: Roberts, Chris
 APPLICANT: Van 't Veer, Laura
 APPLICANT: Van de Vijver, Marc
 APPLICANT: Bernards, Rene
 TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
 FILE REFERENCE: 9301-175-999
 CURRENT APPLICATION NUMBER: US/10/172,118
 CURRENT FILING DATE: 2002-06-14
 PRIOR APPLICATION NUMBER: 60/380,770
 PRIOR FILING DATE: 2002-05-14
 NUMBER OF SEQ ID NOS: 2699
 SEQ ID NO 1685
 LENGTH: 2487
 TYPE: DNA
 ORGANISM: Homo sapiens
 PUBLICATION INFORMATION:
 DATABASE ACCESSION NUMBER: NM_017585
 DATABASE ENTRY DATE: 2001-06-18
 US-10-172-118-1685

Alignment Scores: Length: 2487
 Pred. No.: 9,86e-88 Matches: 227
 Score: 958.50 Conservative: 78
 Percent Similarity: 59.22% Mismatches: 164
 Best Local Similarity: 44.08% Indels: 47
 Query Match: 39.01% Gaps: 12
 DB:

US-09-886-954A-1 (1-477) x US-10-172-118-1685 (1-2487)

QY 3 ProGlu--AspProGluGluThrGlnProLeuLeuGly-----14
 Db 20 CCCGAGAGAGACCCGGCCATGCGAGCCCTGCTGGAGCCCGGGCCGACTACGAC 79
 QY 15 -----ProProGlyGlySerAlaProArgGlyArg-----24
 Db 80 ACCITCCCAGAGAGCCGCCCTCGCCAGGGGACAGGGCGGGGTCGGGACCCCTGCAG 139
 QY 25 -----ArgValPheLeuAlaPheAlaAlaAlaLeuGlyProLeuSerPheGlyPhe 42
 Db 140 AACAAAGGGGTTCCTGGCCACCCTTCGCCGACATGCTCGGCAATTCAGCTTTGGGTAT 199
 QY 43 AlaLeuGlyTySerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaPro 62
 Db 200 GCCCTGGTCTACACATCCCTGTATCCAGCCCTCGAGGCTCCTTGGATCTGAGCTG 259
 QY 63 ArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValThrLeuGlyAlaAlaAla 82
 Db 260 CATCTGACCAAAATCCAGGCATCCTGGTTCGCGTTCCTGGTTCGCGGAGCCGCC 319
 QY 83 GlyGlyValLeuLeuGlyTrpLeuValAspArgAlaGlyArgGlyLeuSerLeuLeu 102
 Db 320 GGAGCCCTGAGTGGCCATGATCTCAACGACCTCCCTGGCCCGGAGCTGAGCATCATGTC 379
 QY 103 CysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMet 122
 Db 380 TCAGCTGTCCCTCGCGCCGCTATGCTGCTATGCGCGGGTTCGCGGAGCCGCTCTGGATG 439
 QY 123 LeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaPro 142
 Db 440 CTGCTCTCGGAAGGACGCTGACGGGCTTCGCCGGGGGCTTCACAGCTGCCCTCATCCCG 499

Db 432 CTGCTGCTCGAAGACGCTGACGGGCTTCGCCGGGGGCTCACAGCTGCGCTGCATCCCG 491
 QY 143 ValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGln 162
 Db 492 GTGTACTGCTGTGAGATTGCTCCCGCAGCGCTTCGTTGGGGCTCTGGGGGCGCACCCAC 551
 QY 163 LeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGluTrpArg 182
 Db 552 CTCATGGCAGTGTTCGATCCCTGTCCTTACCGCCCTTGGCCCTGCTGCGCGTGGCG 611
 QY 183 TrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuLeuMetCysPheMet 202
 Db 612 TGGCTGGCTGTGGCCGGGAGCGGCTGCTCATCATGATCATGCTGCTGCTCATG 671
 QY 203 ProGluThrProArgPheLeuThrGlnHisArgArgGlnGluAlaMetAlaAlaLeu 222
 Db 672 CCCAACTCCCGGCTTCCTGCTCTCTCGGGCAGGAGAGAGCCCTCGGGGCGTGTG 731
 QY 223 ArgPheLeuTrpGlySerGluGlnGly-----TrpGluAspProProIleGly-----238
 Db 732 CCTGTGCTGGTGGAGCGAGCTCGATGTCCTGAGTTCGAGCAGATCCAGGACAAC 791
 QY 239 ---AlaGluGlnSerPheHisLeu-----AlaLeuLeuArgGlnProGlyIleTyrIys 255
 Db 792 GTCGGAGACAGACAGCAGCGAGTATCGTGGGTGAGGCACGGCCCGCCACGCTGCGCGG 851
 QY 256 ProPheIleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaVal 275
 Db 852 CCCATCACCTGCGCTGTGTGATGCGCCCTCGCAGCAGCTGACGGGCATCAGCCCATC 911
 QY 276 MetPheTyrAlaGluThrIlePheGluGluAlaLysPhe-----LysAspSer 291
 Db 912 CTGCTACTGCTGAGCAGCTTCGACAGCAGCCGCTGCTGCTGCTGCTGCTGCTGCTG 971
 QY 292 SerLeuAlaSerValValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeu 311
 Db 972 -----GCAGCAATGTTGGGGCGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1022
 QY 312 IleMetAspArgAlaGlyArgArgLeuLeuLeuValLeuSerGlyValValMetValPhe 331
 Db 1023 ACCATGGACCTCGCAGCCCGCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1082
 QY 332 SerThrSerAlaPheGlyAlaTyrPheLysLeuThrGlnGlyProGlyAsnSerSer 351
 Db 1083 GCCAACTGACTCTGGGGCTGTACATCCATTT-----GCCCGCAGGCTCTGAGC 1133
 QY 352 HisValAlaIleSerAlaProValSer-----AlaGlnProValAsp 365
 Db 1134 CCCAACGACTGGGGCCCGAAGCGAGTCCCTGGGGGACTTGGCCGACCCCTGGCA 1193
 QY 366 AlaSerValGly---LeuAlaTrpLeuAlaValGlySerMetCysLeuPheIleAlaGly 384
 Db 1194 GCACCCGTGGTACTACCTCCCTGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1253
 QY 385 PheAlaValGlyTrpGlyProIleProTrpLeuMetSerGluIlePheProLeuHis 404
 Db 1254 TACGCGGGTGGTGGGTCCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1313
 QY 405 ValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPheLeuVal 424
 Db 1314 GCCCGTGGCTGGCTCAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1373
 QY 425 ThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTrpLeuAla 444
 Db 1374 ACCAAGTCTCTCCCAAGTGTG-----GAGCACCTTGGCCCTCCAGTGCCTTCTTCTT 1432
 QY 445 -SerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluThrLysG 464
 Db 1433 CGGGCCATCTCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1492
 QY 464 yLysThrLeuGluGlnIleThrAlaHisPheGlu---GlyArg 477

QY 143 ValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGln 162
 Db 500 GTTACCGTGTGAGATGCTCCCGGCGGTTCTGGGGCTCTGGGGGCCACACCCAG 559
 QY 163 LeuMetValValGlyIleLeuLeuAlaTyrTrpValLeuGluTyrArg 182
 Db 560 CTATGGCAGTTCGATCCCTTCCTCTACGCCCTTGGCCCTCTGCTGCGGTGGCCG 619
 QY 183 TrpLeuAlaValLeuGlyCysValProSerLeuMetLeuLeuLeuMetCysPheMet 202
 Db 620 TGGCTGGCTGTGCCGGGGAGCGCTGTGCTCATCATGATCTGCTGCTCAGCTTCATG 679
 QY 203 ProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGlnAlaMetAlaAlaLeu 222
 Db 680 CCCAACTCGCGGCTTCTGCTCTCTGGGGCAGGGACGAGAGGCGCTTGGGGCGCTG 739
 QY 223 ArgPheLeuTrpGlySerGluGlnGly-----TrpGluaspProIleGly----- 238
 Db 740 GCTTGGCTGGTGGAGCGAGCTCGATGTCCTCTGGGAGTTCGAGCAGATCCAGGACAAC 799
 QY 239 ---AlaGluGlnSerPheHisLeu-----AlaLeuLeuArgGlnProGlyIleTyrLys 255
 Db 800 GTCCGGAGACAGACGCGGATCTGGGCTGAGGACGGGCCCCACACGCTGTCGGCG 859
 QY 256 PropheIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAenAlaVal 275
 Db 860 CCCATCAGCGCTTGTGTGCGCTCTCTGAGCAGCTGACGGGCATCAGCCCATC 919
 QY 276 MetPheTyrAlaGluThrIlePheGluGluAlaLysPhe-----LysAspSer 291
 Db 920 CTGGTCTACCTGAGTCCAFCTTCAGACAGCAGCCGCTGCTGCTGCCCCCAAGGACGAC 979
 QY 292 SerLeuAlaSerValValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeu 311
 Db 980 -----GCACCATCTGTGGGGCGTGGCTCTCTGCTGAGCGCCATCATGTTGCT 1030
 QY 312 IleMetAspArgAlaGlyArgLeuLeuValLeuSerGlyValValMetValPhe 331
 Db 1031 ACCATGGACTCGCAGCGCGCAAGTGTCTTCTGCTCAGCGCCATCATGTTGCT 1090
 QY 332 SerThrSerAlaPheGlyAlaTyrPheLysLeuThrGlnGlyProGlyAsnSerSer 351
 Db 1091 GCCAACCTGACTCTGGGGGTGTACATCCACTTT-----GGCCCCAGGCTCTGAGC 1141
 QY 352 HisValAlaIleSerAlaProValSer-----AlaGlnProValAsp 365
 Db 1142 CCCAACGACTCGGGCGCTGGAAAGCGAGTCTGGGGGACTTGGCGGAGCCCTGGCA 1201
 QY 366 AlaSerValGly---LeuAlaTrpLeuAlaValGlySerMetCysLeuPheIleAlaGly 384
 Db 1202 GCACCCGCTGGCTTACCTCACCCCTGGTCCCTGCTGGCCACCACATGCTTTCATCANGGC 1261
 QY 385 PheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePheProLeuHis 404
 Db 1262 TAGCCGCTGGGCTGGGCTCCCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1321
 QY 405 ValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPheLeuVal 424
 Db 1322 GCCCGTGGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1381
 QY 425 ThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTrpLeuAla 444
 Db 1382 ACCAAGTCTTCTGCGCAGTGGT-GAGCACCTTGGGCTCCAGGTGCTTCTTCTTCTT 1440
 QY 445 -SerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluThrLysGly 464
 Db 1441 CGCGGCCATCTGCTGGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 QY 464 YLysThrLeuGluGlnIleThrAlaHisPheGlu---GlyArg 477
 Db 1501 ACGGTCCTGGAGCAGATCGATGCTTCTTCCGATGGGGAGA 1543

US-10-342-887-1685
 ; Sequence 1685, Application US/10342887
 ; Publication No. US20040058340A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dai, Hongyue
 ; APPLICANT: He, Yudong
 ; APPLICANT: Linsley, Peter S.
 ; APPLICANT: Mao, Mao
 ; APPLICANT: Roberts, Christopher J.
 ; APPLICANT: Van 't Veer, Laura Johanna
 ; APPLICANT: Van de Vijver, Marc J.
 ; APPLICANT: Bernards, Rene
 ; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
 ; FILE REFERENCE: 9301-188-999
 ; CURRENT APPLICATION NUMBER: US/10/342,887
 ; CURRENT FILING DATE: 2003-01-15
 ; PRIOR APPLICATION NUMBER: 60/298,918
 ; PRIOR FILING DATE: 2001-06-18
 ; PRIOR APPLICATION NUMBER: 60/380,710
 ; PRIOR FILING DATE: 2002-05-14
 ; PRIOR APPLICATION NUMBER: 10/172,118
 ; PRIOR FILING DATE: 2002-06-14
 ; NUMBER OF SEQ ID NOS: 2699
 ; SEQ ID NO 1685
 ; LENGTH: 2487
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-342-887-1685
 Alignment Scores: Length: 2487
 Pred. No.: 9,86e-88 Matches: 227
 Score: 958.50
 Percent Similarity: 59.22% Conservative: 78
 Best Local Similarity: 44.08% Mismatches: 164
 Query Match: 39.01% Indels: 47
 DB: 17 Gaps: 12
 US-09-886-954A-1 (1-477) x US-10-342-887-1685 (1-2487)
 QY 3 ProGlu---AspProGluGluThrGlnProLeuLeuGly----- 14
 Db 20 CCCGAGAGAGACCCCGCCATCGAGCGCTGCTGGGAGCCGAGGGCCCGGACTACGAC 79
 QY 15 -----ProProGlyGlySerAlaProArgGlyArg----- 24
 Db 80 ACCTTCCCCGAGAGCCGCCCTGCGCCAGGGGACAGGGCGGGTCCGGGACCTGCGAG 139
 QY 25 -----ArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPhe 42
 Db 140 AACAAAAGGGTCTCTGGCCACCTTCCCGCAGTGTCTCGCAATTTTCAGCTTTGGGTAT 199
 QY 43 AlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProAlaPro 62
 Db 200 GCCCTGGTCTACACATCCCTGTCATCCAGCCCTGGAGCGCTCTTGGATCCTGACCTG 259
 QY 63 ArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAla 82
 Db 260 CATCTGACCMAATCCAGGCATCTTGTGGTTCGTGTTTCCCTCCCTGGGAGCAGCGGCC 319
 QY 83 GlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeu 102
 Db 320 GGAGCCCTGAGTGCATGATCTCAACGACCTCTCCGGCCGGAAGCTGAGCATCANGTTC 379
 QY 103 CysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMet 122
 Db 380 TCAGCTGTGCCCTCGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 439
 QY 123 LeuLeuGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaPro 142
 Db 440 CTGCTGCTCGGAAGACCGCTGACGGCTTCCCGGGGGGCTCACAGCTCCCTGATCCCG 499
 QY 143 ValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGln 162

Publication No. US20040171823A1
 GENERAL INFORMATION:
 APPLICANT: Bristol-Myers Squibb Company
 TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF- κ B
 TITLE OF INVENTION: PATHWAY
 FILE REFERENCE: D0284 NP
 CURRENT APPLICATION NUMBER: US/10/755,889
 CURRENT FILING DATE: 2004-01-13
 PRIOR APPLICATION NUMBER: U.S. 60/440,068
 PRIOR FILING DATE: 2003-01-14
 PRIOR APPLICATION NUMBER: U.S. 60/469,757
 PRIOR FILING DATE: 2003-05-12
 NUMBER OF SEQ ID NOS: 823
 SOFTWARE: Patent in version 3.2
 SEQ ID NO 109
 LENGTH: 2487
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-755-889-109

Alignment Scores: 9.86e-88 Length: 2487
 Pred. No.: 3 Matches: 227
 Score: 958.50 Conservative: 78
 Percent Similarity: 59.22% Mismatches: 164
 Best Local Similarity: 44.08% Indels: 47
 Query Match: 39.01% Gaps: 12
 DB:

US-09-886-954A-1 (1-477) x US-10-755-889-109 (1-2487)

QY 3 ProGlu---AspProGluGluThrGlnProLeuLeuGly----- 14
 Db 20 CCCGAGAGACCCGGCCATCGAGAGCCGCTGGGAGCCGAGCCGAGCCGCGACTACGAC 79
 QY 15 -----ProProGlyGlySerAlaProArgGlyArg----- 24
 Db 80 ACCTTCCCGAGAGAGCCGCCCCCGCTCCGAGGGGACAGGGCGGGTCCGGACCCCTCGCAG 139
 QY 25 -----ArgValPheLeuAlaAlaPheAlaAlaLeuLeuGlyProLeuSerPheGlyPhe 42
 Db 140 AACAAAGGGGTTCCTGGCCACCCTTCGCGCAGTGTCCGCAATTCAGCTTTGGGTAT 199
 QY 43 AlaLeuGlyTyrSerProAlaIleProSerLeuGlnArgAlaAlaProAlaPro 62
 Db 200 GCCTGTGTTACACATCCCTGTCTCCAGCCCTGGAGCGCTCTTGGATCTCTGACCTG 259
 QY 63 ArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAla 82
 Db 260 CATCTGACCAAATCCAGGCATCTGTGTTGGTTCCTGTTTCCCTCCCTGGGAGCAGCGCC 319
 QY 83 GlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeu 102
 Db 320 GGAGCCCTGAGTGCATGATCTCAACGACCTCTCTGGCCGCGAGCTGAGCATCATGTC 379
 QY 103 CysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMet 122
 Db 380 TCAGCTGTCCGCTCGCGCGCGCTATGCGCTCATGGCGGTGGCGGTCACCGCTCTGGATG 439
 QY 123 LeuLeuGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaAlaSerLeuValAlaPro 142
 Db 440 CTGCTGCTCGAAGACGCTGACGGGCTTCGCGGGGGGCTCACAGCTCCCTGCATCCG 499
 QY 143 ValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGln 162
 Db 500 GGTGACGCTGTGAGATTGCTCCCGAGCGTTCGTGGGGCTCTGGGGCCACACCCGAC 559
 QY 163 LeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGluTrpArg 182
 Db 560 CTCATGGCAGTGTTCGGATCCCTGTCCTACGCGCTTGGCCCTCTCTGCTGCGGGGCG 619
 QY 183 TrpLeuAlaValLeuGlyCysValProSerLeuMetLeuLeuMetCysPheMet 202
 Db 620 TGGCTGGCTGTGGCGGGAGCGGCTGTGCTCATCATGATCCTGCTCATGCTCATGCTCATG 679

500 GTGTACGTTCTGAGATTGCTCCCGAGCGGTTCTGGGGTCTGGGGCCACACCCCGAG 559
 QY 163 LeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGluTrpArg 182
 Db 560 CTCATGGCAGTGTTCGGATCCCTGTCCCTACAGCCCTTGGCTCTCTGCTGCGCGC 619
 QY 183 TrpLeuAlaValLeuGlyCysValProSerLeuMetLeuLeuMetCysPheMet 202
 Db 620 TGGCTGGCTGTGGCGGGAGCGGCTGTGCTCATCATGATCCTGCTCATGCTTCATG 679
 QY 203 ProGluThrProArgPheLeuThrGlnIleSArgArgGlnGlnAlaMetAlaAlaLeu 222
 Db 680 CCAACTGCGCCGCTTCTGCTCTCGGGCAGGAGAGAGGCGCCCTCGGGCGCTG 739
 QY 223 ArgPheLeuTrpGlySerGlnGly-----TrpGluAspProPheIleGly----- 238
 Db 740 GCCTGGCTCGGGAGCGAGCTGCATGTCCTGAGGATTCGAGCAGATCCAGGACAAC 799
 QY 239 ---AlaGlnSerPheHisLeu-----AlaLeuLeuArgGlnProGlyIleTyrLys 255
 Db 800 GTCCGAGACAGACGAGCGAGTATCGTGGCTGAGGACCGGCCCCACAGCTGTGCCG 859
 QY 256 ProPheIleIleGlyValSerLeuMetAlaPheGlnLeuSerGlyValAsnAlaVal 275
 Db 860 CCATCACCGTGGCTGTGATGGCCCTCTGAGCAGCTGACGGGATCAGGCCCATC 919
 QY 276 MetPheTyrAlaGluThrIlePheGluAlaLysPhe-----LysAspSer 291
 Db 920 CTGGTCTACGTCAGTCCATCTTCAGACAGCCGCTGCTCTGCTGCCCCCAAGCAGCAG 979
 QY 292 SerLeuAlaSerValValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeu 311
 Db 980 -----GCAECACCTGTGGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1030
 QY 312 IleMetAspArgAlaGlyArgLeuLeuValLeuSerGlyValValMetValPhe 331
 Db 1031 ACCATGACCTCGAGCGCGCAAGGTGCTGCTTCTGCTCAGCGCCATCATGTTGCT 1090
 QY 332 SerThrSerAlaPheGlyAlaTyrPheLysLeuThrGlnGlyProGlyAsnSerSer 351
 Db 1091 GCCAACCTGACTCTGGGGTGTACATCCACTTT-----GGCCCCAGGCGCTCGAGC 1141
 QY 352 HisValAlaIleSerAlaProValSer-----AlaGlnProValAsp 365
 Db 1142 CCACAGCAGCTGCGGGCTGGAACGAGTCTGGGGGACTTGGCGGCGCCCTGGCA 1201
 QY 366 AlaSerValGly---LeuAlaTrpLeuAlaValGlySerMetCysLeuPheIleAlaGly 384
 Db 1202 GCATCCGCTGGTACCTCACCCCTGGGCGCCCTGCTGGCCACCATGCTTTCATCATGG 1261
 QY 385 PheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePheProLeuHis 404
 Db 1262 TAGCCGCTGGCTGGGGTCCCATCCCTGCTCATGCTGAGTCTGAGGCTCTGCCCCCTG 1321
 QY 405 ValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPheLeuVal 424
 Db 1322 GCCCGTGGGCTGAGGCTCAGGGCTCTGGGCTGGCCAGTGGCTCACCCGCTTGGTCTC 1381
 QY 425 ThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTrpLeuAla 444
 Db 1382 ACCAAGTCTTCTGCGCAGTGGT-GAGCAGCTTGGGCTCCAGGTGCTTCTTCTTCT 1440
 QY 445 -SerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluThrLysG 464
 Db 1441 CGCGGCCATCTGTTGGTGAAGCTGGGTTCACAGGCTGCTGTGGTGGCCGAGCAAGGG 1500
 QY 464 LysThrLeuGluGlnIleThrAlaHisPheGlu---GlyArg 477
 Db 1501 ACGGTCTCGAGCAGATCGAGTCTTTCGCGATGGGAGA 1543

RESULT 10
 US-10-755-889-109
 ; Sequence 109, Application US/10755889

QY 203 ProGluThrProArgPheLeuLeuThrGlnHisArgGlnGluAlaMetAlaLeu 222
 Db 680 CCCAACTCGCGCGCTTCTCTCTCGGGGAGGAGCGAGGAGCGCCCTCGCGGGCTG 739
 QY 223 ArgPheLeuTrpGlySerGluGlnGly-----TipGluAspProFroIleGly----- 238
 Db 740 GCCTGGCTGGTGGGACGAGCGTGTGTCCTACTGGGAGTTCAGGAGATCCAGGACAAC 799
 QY 239 ---AlaGluGlnSerPheHisLeu-----AlaLeuLeuArgGlnProGlyIleTrpLys 255
 Db 800 GTCCGGAGACAGAGCCGAGTATCGTGGCTGAGGACGGCCCGCCACAGTGTCCCG 859
 QY 256 ProPheIleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaVal 275
 Db 860 CCATCACTCGCTGGT 919
 QY 276 MetPheTyrAlaGluThrIlePheGluGluAlaLysPhe-----LysAspSer 291
 Db 920 CTGGTCTACTCGTCCAGTCCATCTTCAGACAGCAGCCGCTGTCTGTGCCCCCAAGGACGAC 979
 QY 292 SerLeuAlaSerValValValGlyValIleGlnValLeuPheThrAlaValAlaLeu 311
 Db 980 ---i-----GCACCATGTGGGGCGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1030
 QY 312 IleMetAspArgAlaGlyArgArgLeuLeuValLeuSerGlyValValMetValPhe 331
 Db 1031 ACCATGGACTCGCAGCGCCGCAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1090
 QY 332 SerThrSerAlaPheGlyAlaTyrPheLysLeuThrGlnGlyGlyProGlyAsnSerSer 351
 Db 1091 GCCAACCTGACTCTGGGGGTGTACATCACTTT-----GGCCCCAGGCTCTTGAGC 1141
 QY 352 HisValAlaIleSerAlaProValSer-----AlaGlnProValAsp 365
 Db 1142 CCCAACGACTCGGGCCCTGGAAAGAGTCTGGGGGACTTGGCGGAGCCCTTGCCA 1201
 QY 366 AldSerValGly---LeuAlaTrpLeuAlaValGlySerMetCysLeuPheIleAlaGly 384
 Db 1202 GCACCCGCTGGCTACTCACTCCCTGGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1261
 QY 385 PheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePheProLeuHis 404
 Db 1262 TAGCCGCTGGTGGGCTCCCACTCACTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1321
 QY 405 ValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPheLeuVal 424
 Db 1322 GCCCGTGGCTGGCCCTCAGGCTCTGGTGGCCAGTGGCTCACCCTTGGCTCTC 1381
 QY 425 ThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTrpLeuAla 444
 Db 1382 ACCAAGTCTCTCTGGCAGTGTG-GAGCACCTTGGGCTCCAGGTGGCTTCTCTCTCTT 1440
 QY 445 -SerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluThrLysG1 464
 Db 1441 CGCGGCACTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1500
 QY 464 YLysThrLeuGlnIleThrAlaHisPheGlu---GlyArg 477
 Db 1501 ACGTCCCTGGAGCAGATCGAGTCTTCTTCCGCAATGGGGAGA 1543

;/ CURRENT APPLICATION NUMBER: US/09/814,353
 ;/ CURRENT FILING DATE: 2001-03-21
 ;/ PRIOR APPLICATION NUMBER: US 60/191,031
 ;/ PRIOR FILING DATE: 2000-03-21
 ;/ PRIOR APPLICATION NUMBER: US 60/207,124
 ;/ PRIOR FILING DATE: 2000-05-25
 ;/ PRIOR APPLICATION NUMBER: US 60/211,940
 ;/ PRIOR FILING DATE: 2000-06-15
 ;/ PRIOR APPLICATION NUMBER: US 60/216,820
 ;/ PRIOR FILING DATE: 2000-07-07
 ;/ PRIOR APPLICATION NUMBER: US 60/220,661
 ;/ PRIOR FILING DATE: 2000-07-25
 ;/ PRIOR APPLICATION NUMBER: US 60/257,672
 ;/ PRIOR FILING DATE: 2000-12-21
 ;/ NUMBER OF SEQ ID NOS: 22037
 ;/ SOFTWARE: FASTSEQ for Windows Version 4.0
 ;/ SEQ ID NO 19608
 ;/ LENGTH: 2696
 ;/ TYPE: DNA
 ;/ ORGANISM: Homo sapiens
 ;/ FEATURE:
 ;/ NAME/KEY: misc_feature
 ;/ LOCATION: 1..2525..2526..2527..2528..2529..2530..2531..2532..2533..
 ;/ LOCATION: 2534..2535..2536..2537..2538..2539..2540..2541..2542..2543..
 ;/ LOCATION: 2544..2545..2546..2547..2548..2549..2550..2551..2552..2553..
 ;/ LOCATION: 2554..2555..2556..2557..2558..2559..2560..2561..2562
 ;/ OTHER INFORMATION: n = A,T,C or G
 ;/ FEATURE:
 ;/ NAME/KEY: misc_feature
 ;/ LOCATION: 2563..2564..2565..2566..2567..2568..2569..2570..2571..2572..
 ;/ LOCATION: 2573..2574..2575..2576..2577..2578..2579..2580..2581..2582..
 ;/ LOCATION: 2583..2584..2585..2586..2587..2588..2589..2590..2591..2592..
 ;/ LOCATION: 2593..2594..2595..2596..2597..2598..2599..2600..2601
 ;/ OTHER INFORMATION: n = A,T,C or G
 ;/ FEATURE:
 ;/ NAME/KEY: misc_feature
 ;/ LOCATION: 2602..2603..2604..2605..2606..2607..2608..2609..2610..2611..
 ;/ LOCATION: 2612..2613..2614..2615..2616..2617..2618..2619..2620..2621..
 ;/ LOCATION: 2622..2623..2624..2625..2626..2627..2628..2629..2630..2631..
 ;/ LOCATION: 2632..2633..2634..2635..2636..2637..2638..2639..2640..2641..
 ;/ LOCATION: 2642..2643..2644..2645..2646..2647..2648..2649..2650..2651..
 ;/ LOCATION: 2652..2653..2654..2655..2656..2657..2658..2659..2660..2661
 ;/ OTHER INFORMATION: n = A,T,C or G
 ;/ FEATURE:
 ;/ NAME/KEY: misc_feature
 ;/ LOCATION: 2695..2696
 ;/ OTHER INFORMATION: n = A,T,C or G
 ;/ US-09-814-353-19608

Alignment Scores:
 Pred. No.: 1..1e-87 Length: 2696
 Score: 958.50 Matches: 227
 Percent Similarity: 59.23% Conservatave: 78
 Best Local Similarity: 44.08% Mismatches: 164
 Query Match: 39.01% Indels: 47
 DB: 10 Gaps: 12

US-09-886-954A-1 (1-477) x US-09-814-353-19608 (1-2696)
 QY 3 ProGlu---AppProGluLuThrGlnProLeuLeuGly----- 14
 Db 55 CCCGAGAGAGACCCTGGCCATCGAGGCGCTGTCTGGGAGCGCGGGCCCGGACTACGAC 114
 QY 15 -----ProProGlyGlySerAlaProArgGlyArg----- 24
 Db 115 ACCTTCCTCCCGAAGCCCGCCCGCTCCGAGGGAGACAGGGCCGGTCTGGGCTGCAG 174
 QY 25 -----ArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPhe 42
 Db 175 AACAAAAGGTTGTTCTCTGCCACCTTCGCGGAGTCTCGGCAATTTTCAGTTTGGTAT 234
 QY 43 AlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProAlaPro 62
 Db 235 GCCTTGGTGTACACATCCCTGTCTATCCAGCCCTCCCTGGAGCGCTCTTGGATCCTGACCTG 294

RESULT 11
 ;/ Sequence 19608 Application US/09814353
 ;/ Publication No. US20030165831A1
 ;/ GENERAL INFORMATION:
 ;/ APPLICANT: Lee, John
 ;/ APPLICANT: Thompson, Pamela
 ;/ APPLICANT: Lillie, James
 ;/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
 ;/ IDENTIFICATION, ASSESSMENT, PREVENTION, AND
 ;/ THERAPY OF OVARIAN CANCER
 ;/ FILE REFERENCE: MRI-006B

QY 63 ArgLeuAspAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAla 82
 Db 295 CArCTgACCAAAATCCAGGATCCTGGTTGGTTCCCGTTCACCCCTGGAGCAGCGCC 354
 QY 83 GlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLeuSerLeuLeu 102
 Db 355 GGAGCGCTGAGTCCATGATCCTCAACAGCCTCCTGGCCGGAAGCTGAGCATCATGTC 414
 QY 103 CysSerValPropPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMet 122
 Db 415 TCAGTGTCCCTCGCGCGCGTATGGCTCATGGCTCATGGGCTGCGCCCTCGATG 474
 QY 123 LeuLeuGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaPro 142
 Db 475 CTGCTGCTCGAAGGACCGTGAAGGCTTCCGCGGGGGCTCACAGCTGCTGCAATCCG 534
 QY 143 ValTyrIleSerGluIleAlaTrpProAlaValArgGlyLeuLeuGlySerCysValGln 162
 Db 535 GNGTAGTGTCTGAGATTGCTCCCCAGCGGTTCGTGGGCTTGGGGCCACACCCAG 594
 QY 163 LeuMetValValGlyIleLeuAlaTyrLeuAlaGlyTrpValLeuGluTrpArg 182
 Db 595 CTCATGGCAGTGTTCGGATCCCTGCTCAAGCCTTGGCCTTGGCCTGCTGCTGGG 654
 QY 183 TrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPheMet 202
 Db 655 TGGTGGCTGTGGCGGGGAGCGCTGTGCTCATCAATGCTTCAGTCTGCTCAGTTCATG 714
 QY 203 ProGluTrpProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaAlaLeu 222
 Db 715 CCCAACTCCCGCGCTTCTCTCGGGGAGCGGACGAAGAGCGCCCTGGGGCGCTG 774
 QY 223 ArgPheLeuTrpGlySerGluGlnGly-----TrpGluAspProIleGly----- 238
 Db 775 GCCTGGCTCGTGGGACGCTCGATGTCCTCAAGTTCGAGCATCGAGCAAC 834
 QY 239 ---AlaGlnSerPheHisLeu-----AlaLeuLeuArgGlnProGlyIleTyrIlys 255
 Db 835 GTCCGGAGACAGGACCGGATGCTGGTGTGGCTGGACGCGGCGCCACACGCTGTGCCG 894
 QY 256 PropheIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaVal 275
 Db 895 CCATCCCGTGGCTGTGTATGTCCTCGACAGCAGTCAAGGGGATCAAGCCATC 954
 QY 276 MetPheTyrAlaGluThrIlePheGluGluAlaLysPhe-----LysAspSer 291
 Db 955 CTGGTCTAGCTGAGTCCATCTCGACAGCAGCAGCGCTGCTCGTCCCGCCCGCAGGACGAC 1014
 QY 292 SerLeuAlaSerValValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeu 311
 Db 1015 -----GCAGCCATGTTGGGGCGTGGGCTCCTCGTGTGTGATCGCGCCCTC 1065
 QY 312 IleMetAspArgAlaGlyArgArgLeuLeuValLeuLeuSerGlyValValMetValPhe 331
 Db 1066 ACCATGGACTCGCGCGCGAAGTGTCTGCTTCTCGTCAAGCGCCATCATGTGTGT 1125
 QY 332 SerThrSerAlaPheGlyAlaTyrPheLysLeuThrGlnGlyGlyProGlyAsnSerSer 351
 Db 1126 GCAACCTGACTCTGGGGGTGTACATCCATTT-----GGCCCGAGCGCTGTGAGC 1176
 QY 352 HisValAlaIleSerAlaProValSer-----AlaGlnProValAsp 365
 Db 1177 CCCAAGCAGCTCGGGCCCTGGNAAGGAGTCTCGGGGGAGTGTGGCGCAGCCCTGGCA 1236
 QY 366 AlaSerValGly---LeuAlaTrpLeuAlaValGlySerMetCysLeuPheIleAlaGly 384
 Db 1237 GCACCGCTGGCTACCTCACCCTGGTGCCTGCTGTGTCGCCACCATGCTCTCATATGGGC 1296
 QY 385 PheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePheProLeuHis 404
 Db 1297 TAGCCGCTGGCTGGGGTCCCACTACCTGCTGCTCATGTGTGAGGCTCGCCCTCGCT 1356
 QY 405 ValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPheLeuVal 424

Db 1357 GCCGTGGGCTGCAGGCTCTGGTCTGCTGCGCAGTGGCTCACCGCTTCGTCTC 1416
 QY 425 ThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTrpLeuAla 444
 Db 1417 ACCAAGTCTTCTCGCAGTGTGAGCCTTGGTTCACAGGCTGCTGTGTGCCGAGCAAGGG 1475
 QY 445 -SerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluThrLysGly 464
 Db 1476 CGCGCCATCTGCTGGTGGAGCCTTGGTTCACAGGCTGCTGTGTGCCGAGCAAGGG 1535
 QY 464 YLyeThrLeuGluGlnIleThrAlaHisPheGlu---GlyArg 477
 Db 1536 ACGTCTCTCGAGCAGATCGAGTCTTCTCCGACCGGGGAGA 1578
 RESULT 12
 US-10-424-599-132422
 ; Sequence 132422, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
 ; TITLE OF INVENTION: plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 132422
 ; LENGTH: 2202
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(2202)
 ; OTHER INFORMATION: unsure at all n locations
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_90584C.1
 US-10-424-599-132422
 Alignment Scores: Length: 2202
 Score: 5.97e-63 Matches: 164
 Percent Similarity: 716.00 Conservative: 89
 Best Local Similarity: 54.88% Mismatches: 174
 Query Match: 29.14% Indels: 34
 DB: 17 Gaps: 7
 US-09-886-954A-1 (1-477) x US-10-424-599-132422 (1-2202)
 QY 26 ValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeuGly 45
 Db 384 GTCTCTTCTCGCTCCTCATCGTCCCTGGTCCCAATTCATTCAGTTCACGGTGGG 443
 QY 46 TyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaProArgLeuAsp 65
 Db 444 TATTCTTCTCCACCCCAAGGCTATAGTTCGCGATCTA-----AACCTCTCT 491
 QY 66 AspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaGlyGlyVal 85
 Db 492 ATTTCCGAGTTCCTTCTTTGGATCTTGTCTTAATGTGGGGAATGTTGGGAGCTATA 551
 QY 86 LeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeuCysSerVal 105
 Db 552 GCTAGTGTCTAGATAGTGAATACATCGGCGCAAGGCTCATGTATGATTCCTCCGATC 611
 QY 106 PropPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMetLeuLeuGly 125
 Db 612 CCCAATTAATAGGGTGGCTGCTATTTCTTTGCAAAAGATTCCTCGTTTTTTGATATG 671
 QY 126 GlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaProValTyrIle 145

```

672 GGGAGGTTGGTGAAGGTTTGGCGTTGGGAAATATCTTATGTGGTGCCTGTTTATATA 731
QY
146 SerGluIleAlaIleValArgGlyLeuLeuGlySerCysValGlnLeuMetVal 165
Db
732 GCTGAGATGGCACTCAAACTTGAGAGTGGCCCTGGGTGAGTCAACCAAGCTCTCTGT 791
QY
166 ValValGlyIleLeuLeuAlaIleValLeuAlaGlyTrpValLeuGluTrpArgTrpLeuAla 185
Db
792 ACAATTGGCATTATGCTGGCTTATCTGTTGGGTCTTTTGTCACTCACTGGAGCTTCGCA 851
QY
186 ValLeuGlyCysValProSerLeuLeuLeuMetCysPheMetProGluThr 205
Db
852 ATTCTAGGAATTTGGCTGTACAGTAAATCACTGGATATATTTTTCATACCTGAATCC 911
QY
206 ProArgPheLeuLeuThrGlnHisArgGlnGluAlaMetAlaAlaLeuArgPheLeu 225
Db
912 CCCAGATGGTGGCCCAAGATGGGATGATAGATGGTGGAGCTCTTTTTCGAAGTGTTA 971
QY
226 TrpGlySerGluGlnGlyTrpGluAspProProIle----- 237
Db
972 CGAGGATTTGAC-----ACTGATATATCTGTTGMAAGTACATGAATTAAGAGATCT 1022
QY
238 ----GlyAlaGluGlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIle 253
Db
1023 GTGGCTTCAACGGGAAAAAGAGCTGCAATCCGATTTGCGAGATCTCAAGAGGAAAAGATAT 1082
QY
254 TyrLysProPheIleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAan 273
Db
1083 TGGTCCCGTAAATGGTATGGTATGGATTAATCTGTTCTTCCAGCAATATCTGATCAAT 1142
QY
274 AlaValMetPheTyrAlaGluThrIlePheGluGluAlaLysPheLysAspSerLeu 293
Db
1143 GGAATTTGTTCTATTCACTACCACTCTTTGCAAAATGCAGGAATTTTCATCCAGCAAGCT 1202
QY
294 AlaSerValValValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMet 313
Db
1203 GCTPACAGTTGGACTGGAGCCGTTACAGTCATAGCAACTGGAATTTCCACATGGTTGGTG 1262
QY
314 AspArgAlaGlyArgLeuLeuLeuValLeuSerGlyValIleMetValPheSerThr 333
Db
1263 GACAAAGTGGCCGAGGCTGCTTCTAATAATATCTCATCTGTAATGACAGTTAGCCCTT 1322
QY
334 SerAlaPheGlyAlaIleValPheLysLeuThrGlnGlyProGlyAsnSerSerHisVal 353
Db
1323 CTGATTGTTCTATAGCATTTTATCTG---GAGGGGTTGTATCAGAGGATTCACATTTA 1379
QY
354 AlaIleSerAlaProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeu 373
Db
1380 -----TTCAGCAATTTTGGGAATAGTTTCTATTTGTTGACTCGTGGCTATG 1424
QY
374 AlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIlePro 393
Db
1425 GTGATTT-----GGTCTCTCTAGTCTGGACCCATCCCT 1460
QY
394 TrpLeuLeuMetSerGluIlePheProLeuHisValLysGlyValAlaIleThrGlyIleCys 413
Db
1461 TGGCTTATAATGTCAGATCTCCAGTGAATATAAGGGCTTCTGGCAGCATAGCG 1520
QY
414 ValLeuThrAsnTrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMetGlu 433
Db
1521 ACAATGGGAAATTTGGTGGTGGGATCCAGTACTGCTAACTGTTGTTTGAAT 1580
QY
434 ValLeuArgProTyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeu 453
Db
1581 ----TGGAGCAGTGGAGGGACATTTACAATCTACACAGCTCGTAGCTCCCTTACTAGCT 1637
QY
454 PheThrLeuPheCysValProGluThrLysGlyLysThrLeuGluGlnIleThrAlaHis 473
Db
1638 TTTATAGCAATGGGTTCTTGAGACCAGGGNAGAACATTTGGAGAAATTCAGTTCCTCC 1697
QY
474 Phe 474
  
```

```

Db 1698 TTC 1700
RESULT 13
US-10-425-114-26596
; Sequence 26596, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 2003-04-28
; LENGTH: 1797
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4371-017-F11_FLI
US-10-425-114-26596

Alignment Scores:
Pred. No.: 1,248-61 Length: 1797
Score: 702.00 Matches: 170
Percent Similarity: 53.56% Conservative: 78
Best Local Similarity: 36.72% Mismatches: 176
Query Match: 28.57% Indels: 40
DB: 17 Gaps: 6

US-09-886-954A-1 (1-477) x US-10-425-114-26596 (1-1797)
QY 27 PheLeuAlaAlaPheAlaAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeuGlyTyr 46
Db 284 TTCTCTGCACGCTCATCGTGGCTCGGCCCACTCCAGTTCGGATTCACGGGGGCTTC 343
QY 47 SerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaProArg-LeuAspAs 66
Db 344 TCCTCCCGCAGC-----CAGGACGCCATCAT-CCGAGACCTCGACCTCACCTC 390
QY 66 pAlaAlaAlaSerTrpPheGlyAlaValThrLeuGlyAlaAlaAlaGlyGlyValle 86
Db 391 CTCGAGTTCCTCGGTTCGGATCGTCCAAACGTCGCGCCCATGGTTGGGCGGATGC 450
QY 86 uGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeuLeuLeuLeuLeu 106
Db 451 CAGTGGTCAGATGGCCGAGTACATGGGGGCAAGGGTTCATGTGATGCTGCAATTC 510
QY 106 oPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMetLeuLeuGly 126
Db 511 AAACATCATGTTGGTTCCTCCATCTCCCTTTCGAAAGACTCATCGTTCCTTATATGG 570
QY 126 YArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaProValTyrIle 146
Db 571 ACGATGCTCGAGGGGTTGGTGGTGCATCTTATACGGTCCAGTTTACATAGC 630
QY 146 rGluIleAlaIleTrpAlaValArgGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuVal 166
Db 631 AGAATAATCACTCAAAACATGAGAGTCTCTTGGCTCAGTCAATCAGTTATCTCTAAC 690
QY 166 lValGlyIleLeuLeuAlaIleLeuAlaGlyTrpValLeuGluTrpArgTrpLeuAla 186
Db 691 CGTGGTATATTGTTGGCATAATTTGTCGCGCATGTTGTTCTTGGAGGCTCTTCTGT 750
QY 186 lLeuGlyCysValProSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 206
Db 751 AATAGGAATCTTGGCTTCGACTGTGTGATACCTGGCTTATTTCTTCAATCCAGAAATCCC 810
  
```

```

Qy 206 oAqPheLeuLeuThrGlnHisArgGlnGluAlaMetAlaAlaLeuAraGpPheLeuTr 226
Db 811 AAGATGGTTGGCAAGATGAACATGATGATGATTTGGAGACTTCTTACAAAGTTCTGAG 870
Qy 226 pGlySerGluGlnGlyTrpGluAspProIleGlyAlaGlu----- 240
Db 871 GGGATTTGAG-----ACTGACATCAGCGGGAAGTGAATATAAAGAG 915
Qy 241 -----GlnSerPheHisLeuAlaLeuLeuAraGpGlnProGI 252
Db 916 AGCAGTAGCGTCAGCAAAACAAAGGACACAGATCCGTTTTCAAGAAATTAACCAAGAA 975
Qy 252 yIleTyrLysProPheIleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyVa 272
Db 976 ATACCGCACACCCCTAATAACTAGGAATGGCTACTTGTACTGCAACAGCTAAGTGAAT 1035
Qy 272 lAnAlaValMetPheTyrAlaGluThrIlePheGluGluAlaLysPheLysAspSerSe 292
Db 1036 CAATGGAAATATGTTTATGCAAGTAGCATCTTCAAGAGCAGAGCTCACAAACAGTGA 1095
Qy 292 rLeuAlaSerValValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeuI 312
Db 1096 CTTGGTACATGTCATGCTGGTCTATCCAGGTTCTGTCACAGGTTTACAACCTGGTT 1155
Qy 312 eMetAspArgAlaGlyArgLeuLeuValLeuSerGlyValValMetValPheSe 332
Db 1156 ATTAGACAGAGCTGGCCGACGATCTCTTATCATCTTCTGCTGGATGACTCTAAG 1215
Qy 332 rThrSerAlaPheGlyAlaTyrPheLysLeuThrGlnGlyProGlyAsnSerSerHi 352
Db 1216 CCTCTTCAGTTGCTGTTGTTATTTTCTC---AAGGATAGCATTTTCAAGATTTCTCA 1272
Qy 352 sValAlaIleSerAlaProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTr 372
Db 1273 CATGTACTACCC-----TTAAGTAT 1293
Qy 372 pLeuAlaValAlGlySerMetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProII 392
Db 1294 GATCTCTTGGTTGCTTCTGTTGGCTTTGTATTCGGCTTCTCCTTCGGTATGGTGCAT 1353
Qy 392 eProTrpLeuLeuMetSerGluIlePheProLeuHisValLysGlyAlaAlaThrGlyII 412
Db 1354 TCCATGGATCATAAATGTCAGAGATCCTCCGGTATAGTATCAAGAGCTCGCAGGAAGCTT 1413
Qy 412 eCysValLeuThrAsnTrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMe 432
Db 1414 TGCAGCGCTGGCCAACTGGCTTACATFCCTTGGAAATAACAATGACA---GCCAACTTGAT 1470
Qy 432 tGluValLeuAraGpProTyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerVa 452
Db 1471 GCTTACTGGAGTGTGGAGGACCTTTGTCTCTACATGGTCTGAGTGCCTTTCACCCT 1530
Qy 452 lLeuPheThrLeuPheCysValProGluThrLysGlyLysThrLeuGluGlnIleThrAl 472
Db 1531 CGTGTTCCGTCATCTCTTGGTGGCCAGAGCAAAAGGAGAACTCTGAAAGATACAATG 1590
Qy 472 aHisPhe 474
Db 1591 GTCCTTC 1597

```

```

RESULT 14
US-10-437-963-61061
; Sequence 61061, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping

```

```

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 61061
; LENGTH: 2003
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_62530C.1
US-10-437-963-61061

```

```

Alignment Scores: 1.43e-61 Length: 2003
Pred. No.: 702.00 Matches: 170
Score: 53.56% Conservative: 78
Percent Similarity: 36.72% Mismatches: 176
Best Local Similarity: 28.57% Indels: 40
Query Match: 18 Gaps: 6
DB:

```

US-09-886-954A-1 (1-477) x US-10-437-963-61061 (1-2003)

```

Qy 27 PheLeuAlaAlaPheAlaAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeuGlyTyr 46
Db 346 TTCCTCTGCAGCTCATCGTCGCGCTCGGCCCATCCAGTTCGATTCGATTCACGGGGGCTTC 405
Qy 47 SerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaProAraG-LeuAspAs 66
Db 406 TCCTCCCGGAG-----CAGGAGCCCATCAT-CCGAGACCTCGACCTCACCCCT 452
Qy 66 pAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaAlaGlyGlyValLe 86
Db 453 CTCGAGTTCGCGTGTTCGATCGCTGTCCACGTCGGCCCATCGTTGGGGCGATGTC 512
Qy 86 uGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeuLeuLeuCysSerValPr 106
Db 513 CAGTGGTCAGATGCGCCGAGTACATTTGGGGCAAAAGGGTCAATTGATGATTGCTGCAATTC 572
Qy 106 oPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMetLeuLeuGlyGI 126
Db 573 AAACATCATTTGGTGTTCGCTTTCCTTTGCAAAAAGACTCATCGTTCTTTATATGGG 632
Qy 126 YArgLeuLeuThrGlyLeuAlaCysGlyValAlaAlaSerLeuValAlaProValTyrIleSe 146
Db 633 ACGATTGCTCAGGGGTTTGGTGTGGTGTGTCATCTCTTATACGGTGCAGTTTACATGAC 692
Qy 146 rGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGlnLeuMetValVa 166
Db 693 AGAAATATCACCTCAAAACATGAGAGGTGCTTTGGCTCAGTGAATCAGTTATCTGTAAC 752
Qy 166 lValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGluTrpArgTrpLeuAlaVa 186
Db 753 CGTTGGTATATTTGGCATATTTGCTCGGCATGTTTGTTCCTTGGAGGCTTCCTGCTGT 812
Qy 186 lLeuGlyCysValProProSerLeuMetLeuLeuLeuMetCysPheMetProGluThrPr 206
Db 813 AATAGGAATCTTGCCTTGCATCTGTGTGATACCTGGCCTATTTCTTATCAGAAATCCCC 872
Qy 206 oArgPheLeuThrGlnHisArgArgGlnGluAlaMetAlaAlaLeuAraGpPheLeuTr 226
Db 873 AAGATGGTTGGCAAAAGATGAACATGATGATGATTTGGAGACTTCTTACAAAGTTCTGAG 932
Qy 226 pGlySerGluGlnGlyTrpGluAspProIleGlyAlaGlu----- 240
Db 933 GGGATTTGAG-----ACTGACATCAGCGGGAAGTGAATATAAAGAG 977
Qy 241 -----GlnSerPheHisLeuAlaLeuLeuAraGpGlnProGI 252
Db 978 AGCAGTAGCGTCAGCAAAACAAAGGACACAGATCCGTTTTCAAGAAATTAACCAAGAA 1037
Qy 252 yIleTyrLysProPheIleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyVa 272

```

1038 ATACCCACACCCTAATACTAGGAATGGCTACTTGTACTGCAACAGCTAAGTGAAT 1097
 272 lAenAlaValMetPheTyrAlaGluThrIlePheGluGluAlaLysPheLysAspSerSe 292
 1098 CAATGGANATATGTTTATGAGGTAGCATCTCAAGCGAGGCTCACAACAGTGA 1157
 292 rLeuAlaSerValValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeu1 312
 1158 CTTGGCTACATGACCTGCTGCTATCCAGGTTCTTGTACAGGATTACAACCTGGTT 1217
 312 eMetAspArgAlaGlyArgLeuLeuValLeuSerGlyValValMetValPheSe 332
 1218 ATTAGCAGAGCTGGCCGAGGATCTCTTATCATCTCTTCTGCTGGATGACTTAAG 1277
 332 rThrSerAlaPheGlyAlaTyrPheLysLeuThrGlnGlyGlyProGlyAsnSerXh 352
 1278 CCTCCTGGCAGTGGCTGTTGTTATTTTTTCCTC---AAGGATAGCATTTCCAGATTTCTCA 1334
 352 sValAlaIleSerAlaProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTr 372
 1335 CATGTACTACACC-----TTAAGTAT 1355
 372 pLeuAlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValAlaGlyTrpGlyPro1 392
 1356 GAFTCTCTTGGTGGCTTGTGGCTTTTGTAAATGCGCTTCTCTCCGATGGTCCCAT 1415
 392 eProTrpLeuLeuMetSerGluIlePheProLeuHisValLysGlyValAlaThrGly1 412
 1416 TCATGATGATCAATGTCAGAGATCTCCCGGTAGTATCAAGAGTCTCCGAGGAGCTT 1475
 412 eCysValLeuThrAsnTrpLeuMetAlaPheLeuValThrLysGluPheSerLeuMe 432
 1476 TGAGGCTCGCCACTGGCTTACATCTTTGGAATAACAATGACA--GCAAACTTGTAT 1532
 432 tGluValLeuArgProTyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerVa 452
 1533 GCTAGCTGGAGTGTGGAGGACCTTTGTCTCATATGGTGTGGTGGTGGTCTTCCACCT 1592
 452 lLeuPheThrLeuPheCysValProGluThrLysGlyLysThrLeuGluGlnIleThrAl 472
 1593 CGTGTTCATCCCTTGGTGGCCAGAGACAAAGGAGAACTCTCGAAGATACAATG 1652
 472 ahisPhe 474
 1653 GTGCTTC 1659

RESULT 15
 US-10-425-114-24968
 ; Sequence 24968, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jindong
 ; APPLICANT: Zhou, Yinhua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
 ; FILE REFERENCE: Plants and Uses Thereof for Plant Improvement
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 24968
 ; LENGTH: 2049
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: LIB3689-227-G11_F11
 US-10-425-114-24968

Pred. No.: 1.66e-61 Length: 2049
 Score: 701.50 Matches: 173
 Percent Similarity: 51.1% Conservative: 86
 Best Local Similarity: 34.1% Mismatches: 182
 Query Match: 28.5% Indels: 65
 DB: 17 Gaps: 8
 US-09-886-954A-1 (1-477) x US-10-425-114-24968 (1-2049)
 QY 10 GlnProLeuLeuGly-----proProGlyGlySerAlaProArg 22
 Db 390 AAGCCGCTCATCAACACCAGGAGCTGTACCCGATGCCCGCGGGTCCGCTGATGGGC 449
 QY 23 GlyArg-----ArgValPheLeuAlaAlaPheAla----- 32
 Db 450 TCGCGCAATCCAGCCTCATGGAGCGCTTGGGCTCTCGCTTCTCTCTCCGCGAGCTC 509
 QY 33 -----AlaAlaLeuGlyProLeuSerPheGlyPhe 42
 Db 510 GCTATCTCGGCCAGCCTCTGCACGCTTATGTGGCTAGGTCCTCCATCCAGTTCCGTTTC 569
 QY 43 AlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaPro 62
 Db 570 ACATGGGCTACTCTCGCCACG-----CAGACGCCATCATTTGCTGTATCTC 617
 QY 63 ArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaAla 82
 Db 618 GGCCTCTCCCTCTCAGTCTCTCCCTCTCGTTCATTAATCTAGGGGCGATGGTA 677
 QY 83 GlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeuLeu 102
 Db 678 GCGCCATCTCCAGTGGGCACTTCAGAGTATATCGCCGCAAGGGTCTCTCATGATC 737
 QY 103 CysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMet 122
 Db 738 GCTGGATTCANAAATATTTGGTGGCTCGGATATCATTCGAAAGATTCCTCTTTC 797
 QY 123 LeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaPro 142
 Db 798 TTGTTATGGTGGCTGCTAGAGGATTTGGAGTCGGTGTAAATATCGTATACAGTACC 857
 QY 143 ValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGln 162
 Db 858 GTTATATTCAGAAATCGCTCCTCAGGATCAGAGGGAGCTTTGGTTCGTCAATCAG 917
 QY 163 LeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGluTrpArg 182
 Db 918 CTCCTCCGTCAGATTGGTATATTTGCTTGTCTACTTGTGGCATGTTTCTCCCTGGAGA 977
 QY 183 TrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPheMet 202
 Db 978 ATTCCTGCTGTTCTAGGCAATTTACTTGTTCATCCTGATTCCTGGACTGTTCTTTG 1037
 QY 203 ProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaLeu 222
 Db 1038 CCTGAATCCCCAAGTGGTGGCAAAAATGGGAAAGATGGAGGATTTTGAATATTCATTTG 1097
 QY 223 ArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProIleGlyAlaGlu----- 240
 Db 1098 CAAGTTCGAGGAGATTTCAG-----ACAGATATCACACAGAAAGTAAAT 1142
 QY 241 -----GlnSerPheHisLeuAlaLeuLeu 248
 Db 1143 GAAATAAAGAGATCATAGCATCATCAGGAGGAGGACCAACCAATAGGTTCCGCTGATA 1202
 QY 249 ArgGlnProGlyIleTyrLysProPheIleIleGlyValSerLeuMetAlaPheGlnGln 268
 Db 1203 AAACAGAAAGATACAGTGTCTCCCTTGTGATAGGAATCGGTCTCCTCTCCCTGCAGCAG 1262
 QY 269 SerGlyValAsnAlaValMetPheTyrAlaGluThrIlePheGluAlaLysPhe 288
 Db 1263 CTAAGTGGTCAATGGCATTTCTATTTATGTGGCGCATCTTCAAAGCTGCTGTAT 1322

Alignment Scores:

```

QY 289 LysAspSerSerLeuAlaSerValValGlyValIleGlnValLeuPheThrAlaVal 308
Db 1323 ACAACACAGTAATCTAGCAACATTTGGTTTAGGGGCTTTAGGGTGAATGCTACTGGAGTG 1382
QY 309 AlaAlaLeuIleMetAspArgAlaGlyArgArgLeuLeuValLeuSerGlyValVal 328
Db 1383 ACAACCTGGTTGACTGACAAAGCTGGTGAAGGCTTCTTCATTTATTTCCACCACAGGA 1442
QY 329 MetValPheSerThrSerAlaPheGlyAlaTyrPheLysLeuThrGlnGlyProGly 348
Db 1443 ATGGTCATTTACTCTTGTATTTCTGTGTCATTTTGTGAGGACACATAGCTGCT 1502
QY 349 AnSerSerHisValAlaIleSerAlaProValSerAlaGlnProValAspAlaSerVal 368
Db 1503 GGTTCGCACCTTACTCTGTA----- 1523
QY 369 GlyLeuAlaTrpLeuAlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValGly 388
Db 1524 ---ATGAGTATGCTTTCACCTGGCTGGACTTGTGGCATTGTGATTTGCAATTTCTCTGGC 1580
QY 389 TrpGlyProIleProTrpLeuLeuMetSerGluIlePheProLeuHisValLysGlyVal 408
Db 1581 TTGGGAGCGATTCCGTGGATCATATGCTGAGATCCTTCTGTTTAAACATCAAGAGCCTT 1640
QY 409 AlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPheLeuValThrLysGluPhe 428
Db 1641 GCTGGAAGTGTGGCGACCCTGGCAACTGGCTGACAGCATGGCCATTACAATGAGC--- 1697
QY 429 SerSerLeuMetGluValLeuArgProTyrGlyAlaPheTrpLeuAlaSerAlaPheCys 448
Db 1698 GCAGCCTGATGTTGAACTGGAGCAGTGGAGGAAACATTTGCTATCTACGCCGTCGTCT 1757
QY 449 IlePheSerValLeuPheThrLeuPheCysValProGluThrLysGlyLysThrLeuGlu 468
Db 1758 ACCATGCCCTCATTTTCGTGCTTGTGGTGCCTGAGACCAAGGGAAGAACGCTAGAG 1817
QY 469 GlnIleThrAlaHisPhe 474
Db 1818 GAAATCGCCTTCTCATTC 1835

```

Search completed: February 25, 2005, 23:02:57
Job time : 708 secs

10	1269	51.6	1004	9	AY414181	AY414181 Pan trogl
11	1267	51.6	875	6	CD557801	CD557801 AGENCOURT
12	1241.5	50.5	880	4	BM044230	BM044230 603621486
13	1238	50.4	928	4	EG478000	EG478000 602522606
14	1233	50.2	777	4	EG717034	EG717034 602689093
15	1230.5	50.1	889	4	BI757409	BI757409 603029344
16	1222.5	49.8	879	4	EG771736	EG771736 602720404
17	1193.5	48.6	838	4	EG281777	EG281777 602402292
18	1183.5	48.2	801	4	BI114765	BI114765 602861314
19	1120.5	45.7	671	7	CN306077	CN306077 170006000
20	1120.5	45.6	894	5	BQ919165	BQ919165 AGENCOURT
21	1119	45.5	999	4	BI916408	BI916408 603178364
22	1116	45.4	854	4	BI828656	BI828656 603078566
23	1076	43.8	813	4	BI831908	BI831908 603078776
24	1074.5	43.7	1002	7	BI828949	BI828949 603078419
25	1065.5	43.4	736	7	CK455384	CK455384 915964 MA
26	1059	43.1	833	4	BI826390	BI826390 603076257
27	1055	42.9	700	5	BQ109572	BQ109572 imagec7
28	1036.5	42.2	934	2	BE910478	BE910478 601501005
29	1030.5	41.9	773	5	BQ603775	BQ603775 MI-P-CP1-
30	990	40.3	664	2	BE986058	BE986058 UI-M-CG0P
31	980	39.9	621	2	BE986074	BE986074 UI-M-CG0P
32	973	39.6	920	2	BF308306	BF308306 601890370
33	949	38.6	2125	3	AK089246	AK089246 Mus muscu
34	948	38.6	677	4	EG707107	EG707107 602670234
35	947.5	38.6	684	4	BI907581	BI907581 603065948
36	938	38.2	735	4	BM008943	BM008943 603618748
37	933.5	38.0	2081	3	AK079650	AK079650 Mus muscu
38	931	37.9	994	5	BQ644799	BQ644799 AGENCOURT
39	912.5	37.1	683	6	BY750238	BY750238 BY750238
40	900	36.6	638	2	BB612439	BB612439 BB612439
41	889	36.2	608	4	EG079217	EG079217 H3037C03-
42	888.5	36.2	908	6	CA327128	CA327128 UI-M-FY0-
43	870	35.4	974	5	BQ944156	BQ944156 AGENCOURT
44	869	35.4	594	4	BI402088	BI402088 MI-P-CP0-
45	866	35.2	625	5	BP152931	BP152931 BP152931

ALIGNMENTS

RESULT 1
AK081806
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

AK081806
Mus musculus
Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone: C130078K14 product: solute carrier family 2, (facilitated glucose transporter), member 8, full insert sequence.
AK081806.1 GI:26349396
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2159	87.9	2101	3	AK081806 Mus muscu
2	2123.5	86.4	1954	3	AF289587 Homo sapi
3	1713	69.7	1008	9	AY414180 Homo sapi
4	1490.5	60.7	1098	5	BX395379 BX395379
5	1485.5	60.5	1000	9	AY414182 Mus muscu
6	1379	56.1	1085	4	BM545247 Mus muscu
7	1340	54.5	932	4	EG479842 AGENCOURT
8	1293	52.6	830	4	EG700749 AGENCOURT
9	1269.5	51.7	991	4	BI334832 602998959

OM protein - nucleic search, using frame_plus_p2n model
Run on: February 25, 2005, 18:30:22 ; Search time 3601 Seconds
(without alignments)
5042.115 Million cell updates/sec

Title: US-09-886-954A-1
Perfect score: 2457
Sequence: 1 MTPEDPETQPLPPGGSA.....CVPETKGLQITAHFEGR 477

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool_p/US09886954/runat_23022005_154242_13220/app_query.fasta_1.647
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cgi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09886954.@CGN 1.1 5180 @runat_23022005_154242_13220 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: EST:
1: gb_est1:
2: gb_est2:
3: gb_hic:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gss1:
9: gb_gss2:

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL
 MEDLINE
 PUBMED
 REFERENCES

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 695-690 (2001)
 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II team.

JOURNAL
 REFERENCES
 AUTHORS

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 (bases 1 to 2101)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawaji, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitsu, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

JOURNAL
 REFERENCES
 AUTHORS

Direct Submission
 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL: <http://genome.gsc.riken.jp/>
 URL: <http://fantom.gsc.riken.jp/>
 Location/Qualifiers

JOURNAL
 COMMENT
 FEATURES
 source

```

1..2101
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:C130078K14"
/db_xref="taxon:10090"
/clone="C130078K14"
/tissue_type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days embryo"
61..1494

```

CDS

```

/notes="unnamed protein product; putative solute carrier family 2, (facilitated glucose transporter), member 8 (MGD) [MG1:1860103, GB|NM_019488, evidence: BLASTN, 99%, match=1490]"
/codon_start=1
/protein_id="BAC38338.1"
/db_xref="GI:26349397"
/translations="MSPEDPQETQPLRPPPEARTPRGRVFLASFAALGPLSPFGFAL
GVGSPAI PSLRRTAPALRGNAAQSWGAVVTLGAAGGILGGWLLDRGRKLSILL
CIVPVFTGAVLTAARDVWMLLGRLLTGACGVSALVAPVVISYTAIPAVRGLGSC
VQMLVGTILLAVAGVLEWRWLVAGVFPFVTLMLLMLCYMPEIFRLLTQHVOEA
MAALRFWSSSEGWPEPPVGAERHQGFQALRFRPGYIKFIIIGISLWVFWQLSGLNVAI
MFYANSIFEAKSDSLASVTGTIIQVLFTVAALIMDRGRLLRLLSLLLVWVWFPM
SARGTGFKLQTSLPSNSHVGLVPIAAEFVDVQVGLVLAWLVAGVSMCLIFLAGFAVWNGPI

```

```

PWLMSIFPLHVKGVATGICVLTNWFMAELVTKFESVVMEMLRPLRYPGAFWLTAAFCAL
SVLFTLTVVPETKGRITLQVTAHFEGR"
2080..2085
/polyA_signal
/polyA_site
/notes="putative"
2101
/notes="putative"

```

ORIGIN

Alignment Scores: 4.05e-194 Length: 2101
 Pred. NO.: 2159.00 Matches: 412
 Score: 92.47% Conservative: 30
 Percent Similarity: 86.19% Mismatches: 34
 Best Local Similarity: 92.47% Mismatches: 34
 Query Match: 87.87% Indels: 2
 DB: 3 Gaps: 2

US-09-886-954A-1 (1-477) x AK081806 (1-2101)

QY	1	MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla	20
Db	61	ATGCTCCCGAGGACCCCAGAGACGCAGCGGCTATTGGGCCACC	GGAAAGCCAGGACT 120
QY	21	ProArgGlyArgArgValPheLeuAlaAalApeAlalalaLeuGlyProLeuSerPhe	40
Db	121	CCCCCGCGCCCGGGGTCTCTCGCTCTCTCCGCTCTGGGACCCCTCAGGTTC	180
QY	41	GlyPheAlaLeuGlyTyrSerSerProAlaTleProSerLeuGlnArgAlaAalAalPro	60
Db	181	GGCTTCGGCGTCGCGTACAGCTCCCGGCCATCCAGCCCTTGCGGGCCGCCACCCG	CCGACCCCGG 240
QY	61	AlaProArgLeuAspAlaAalAalSerTrpPheGlyAlaValValThrLeuGlyAla	80
Db	241	GCCCTGGCCCTCGGAGACAAATCGCGCTCTCTGGTTGGGGCGCTCTGGACCTGG	GGCGCT 300
QY	81	AlaAlaglyGlyValLeuGlyGlyTrpLeuValAspArgAlaglyArgLysLeuSerLeu	100
Db	301	GCTGCAGGGGCATATCGGGCGCTTCTCTGGACCGTCCTGACCGTCGAGGGCGAAG	CCTCAGGCCTC 360
QY	101	LeuLeuGlyCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAsp	Val 120
Db	361	TTGCTCTGCACCGTTCCTCTGCTGACTGGCTTGTCTCATCCCGCGCCCGGGATGTG	420
QY	121	TrpMetLeuLeuGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaLaserLeuVal	140
Db	421	TGGATGTCTCGGAGGGCGCGCTCTCACCGCCCTTGGCTCGGAGTGCCCTCATGTG	480
QY	141	AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys	160
Db	481	GCACCGGTTTACATCTCGAAATCGCCCTACCCAGCTGTCGAGGACTGCTCGGCTCCT	540
QY	161	ValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu	180
Db	541	GTCAGCTGATGGTGTTCACCTGGCATCTCTCCCTGGCCATATGGCCAGGCTGGTCT	CCTAGAG 600
QY	181	TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuLeuMetCys	200
Db	601	TGGCGTGGCTGGCGGTGCTGGGGCTGGTGGCCCGCCCGCCACTCAATGCTGCTCAT	660
QY	201	PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGlnAlaMetAla	220
Db	661	TACATGCCAGGACCCCGTCTCTCTCACTCAACACCAGTACAGAGGGCCCAATGGCT	720
QY	221	AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyValaGlu	240
Db	721	GCCCTGCGCTCTCTGTTGGGGCTCTGAGAGGGCTGGGAAAGCCCGCCCTGTTGGGG	CGT 780
QY	241	---GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIle	259
Db	781	CACCAAGGCTTCCAGCTGGCCCTGCTGAGGGCGCCCTGGCATCTCAAGCCCTCATC	ATC 840
QY	260	GlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAla	279
Db	841	GGCATTTCCCTCATGGTCTTCCAGGAGCTGTCCAGGGTCAATGTATCATGTTCTATGCC	900

QY 280 GluThrIlePheGluAlaLysPheLysAspSerSerLeuAlaSerValValGly 299
 Db 901 AACAGCATCTTCGAGGAGCCAAAGTCAAGGACAGCAGCCTGGCTCGCTCACTGGGC 960
 QY 300 ValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArg 319
 Db 961 ATAACTCAGGCTCTGTTCATCTGTGGGGCCCTCATATGGACAGCAGCGGGAAGG 1020

QY 320 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyr 339
 Db 1021 CTGCTCTGGCTTGTGGGTGATCATGATGGTTCAGTATGATGATGATGGTTCCTTGTACCTAC 1080

QY 340 PheLeuLeuThrGlnGlyProGlyAanSerSerHisValAlaIleSerAlaProVal 359
 Db 1081 TTCAAACTCAGCAGGCTCCCCAGCAACTCTCCACGATGGCCCTG--GTGCCCATC 1137

QY 360 SerAlaGlnProValAspAlaSerValGlyLeuAlaTyrLeuAlaValGlySerMetCys 379
 Db 1138 GCGCGGAGCCTGTGGATGTCCAAAGTGGGACTGGCCCTGGCTGTAGGCACGATGTGC 1197

QY 380 LeuPheIleAlaGlyPheAlaValGlyTyrGlyProIleProIleProIleProIleProIle 399
 Db 1198 CTCTTCATCTGTGGCTTTCGGGTTGGGCTGGGACCATCCCTGGCTCCTCATGTGAG 1257

QY 400 IlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeu 419
 Db 1258 ATCTTCCCTCTGCATGTCAAGGTTGGCTACCGGCATCTGTGCTCTACCAACTGGTTC 1317

QY 420 MetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGly 439
 Db 1318 ATGGCTTTCTAGTACCACAAAGATTCAGCAGCGTTCATGGAGATGCTCAGACCCCTACGGT 1377

QY 440 AlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal 459
 Db 1378 GCCTTCTGGCTACCCCTCTCCGCTTCAGTGTCTTCACTACCTGACCGTTGTC 1437

QY 460 ProGluThrLysGlyThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 Db 1438 CCTGAGACTAAAGCAGGACTCTGGAAACAAGTCAACAGCCATTTCGAGGGAGCA 1491

RESULT 2
 AF289587 Homo sapiens (human) 1954 bp mRNA linear HTC 01-JAN-2002
 LOCUS AF289587 Homo sapiens clone pp7425 unknown mRNA.
 DEFINITION AF289587
 ACCESSION AF289587.1 GI:18027377
 VERSION HTC.
 KEYWORDS

SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 AUTHORS Zhang, P.P., Zhou, X.M., Jiang, H.Q., Huang, Y., Qin, W.X., Zhao, X.T.,
 Wan, D.F. and Gu, J.R.
 TITLE Novel human cDNA clones with function of inhibiting cancer cell
 growth
 JOURNAL Unpublished
 REFERENCE Zhang, P.P., Zhou, X.M., Jiang, H.Q., Huang, Y., Qin, W.X., Zhao, X.T.,
 AUTHORS Wan, D.F. and Gu, J.R.
 TITLE Direct Submission
 JOURNAL Submitted (17-JUL-2000) National Laboratory For Oncogenes & Related
 Genes, Shanghai Cancer Institute, 25/Jun 2200 Xie-Tu Road, Shanghai
 200032, P. R. China

FEATURES
 Location/Qualifiers
 1..1954
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="pp7425"
 491..1237
 /codon_start=1

/product="unknown"
 /protein_id="AA155771.1"
 /db_xref="GI:18027378"
 /translation="NVTGILLAVLAGWLVLRWLVAVLGCVPVSLMLLMLCFMPTPR
 ELLTQRRQEAAMALRFPLNGSEQWEDPIGAEQSFHLALLRQPIYKPFILGIVSLMA
 FOQLSGVANMFIYAEIIFEEAFKDDSLASVVVGVIVLFTVAVALIMDRAGRLLLV
 LSGVVMVFTSFAFGAYFKLTQGGPGNSHVVAISPAVPAQVDAVGLAWLAVGSMCLF
 TAGGPOALWLSLACLRLFLHLQPFHFVLP"

ORIGIN

Alignment Scores: 8.69e-191 Length: 1954
 Pred. No.: 2123.50 Matches: 426
 Score: 2123.50
 Percent Similarity: 89.14% Conservative: 1
 Best Local Similarity: 88.94% Mismatches: 1
 Query Match: 86.43% Indels: 51
 DB: 3 Gaps: 1

US-09-886-954A-1 (1-477) x AF289587 (1-1954)

QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
 Db 1 GTGAGCGCCGAGACCCAGAGAAACCCAGCCGCTTCTGGGSCCTTCTGGCGGACGGG 60
 QY 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
 Db 61 CCCCAGCGCGCGCTTCTCTCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTT 120
 QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 Db 121 GCGTTCGCGCTCGCTGACAGCTCCCGCGCCATCCCTAGCTGACAGCGCGCGCGCG 180
 QY 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
 Db 181 GCCCGCGCTGAGAGAGCGCCCGCTCTCGGTTCGGGSCCTCGTGAACCTTGGGTGCC 240
 QY 81 AlaAlaGlyValLeuGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
 Db 241 GCGCGCGCGGAGTGTGGCGCTGGCTGGTGGAGCCGCGCGCGCGCGCGCGCGCGCT 300
 QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaAlaValIleThrAlaAlaGlnAspVal 120
 Db 301 TTGCTGTCTCGCTCGCTTCTGCGCGCTTCTGCGCGCTTCTGCGCGCTTCTGCGCGCT 360
 QY 121 TrpMetLeuLeuGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 Db 361 TGGATGCTGCTGGGCGCGCGCTTCTCACCAGCGCTGGCTGGGCTGGGCTGGGCTGG 420
 QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaVal-ArgGlyLeuLeuGlySerCys 160
 Db 421 GCCCGGCTACATCTCCGMAATCGCTTACCCAGCAGTCCCGGGGCTTCTCGGCTCCTG 480
 QY 160 sValGlnMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGln 180
 Db 481 TGTGAGCTAAATGCTGCTCGCTGCGCATCTCTCGCGCTTCTGCGCGCTTCTGCGCGCT 540
 QY 180 wTrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 Db 541 GGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 600
 QY 200 sPheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGlnAlaMetAl 220
 Db 601 CTTTATGCCGAGACCCCGCGCTTCTGCTGACTCAGACAGCGCGCGCGCGCGCGCGCG 660
 QY 220 aAlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGln 240
 Db 661 CGCCCTGGGTTCTTGTGGGCTCGGAGCAGGCGCTGGGAAGACCCCGCGCGCGCGCG 720
 QY 240 uGlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleGln 260
 Db 721 GCAGAGCTTTCACCTGGGCGCTTCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
 QY 260 yValSerLeuMetAlaPheGlnGlnLeuSerGlyValAenAlaValMetPheTyrAlaGln 280

```

781 CGTCTCCGTGATGATCCAGCCAGTCTCGGGGTCAAGCCCGTCAATGTTTCAATGAGA 840
QY uThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVa 300
Db GACCATCTTTGAAGAGCCAAAGTTCAAGGACAGCAGCCTGGCTCGTCTGGGTGT 900
QY lIleGlnValLeuPheThrAlaValAlaLalaLeuMetAspArgAlaGlyArgAgl 320
Db CATCCAGTCTGTTCAACGTGGGGGTCTCATCATGGACAGAGCAGGGCGGAGGCT 960
QY uLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTrpH 340
Db GCTCTCGTCTGTGTGATGCTGTGTGTGTTTCAGCAGAGTGCCTCGGCGCTACTT 1020
QY eLysLeuThrGlnGlyProGlyAsnSerSerHisValAlaLalaSerAlaProValSe 360
Db CAAGCTGACCCAGGGTGGCCCTGGCACTCTCCGCAGGTGGCCATCTCGCGGCTGTCT 1080
QY rAlaGlnProValAspAlaSerValGlyLeuAlaLalaTrpLeuAlaValGlySerMetCysLe 380
Db TGCACAGCCTGTGTATGCCAGCCTGGGGCTGGCTGGCTGGCGTGGCAGCATGTGCTT 1140
QY uPheIleAlaGly-PheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluI 400
Db CTTTCATCGCGGAGGTC----- 1157
QY lePheProLeuHisValLysGlyValAlaThrGlyLeCysValLeuThrAsnTrpLeuM 420
Db ----- 1157
QY etAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTYrGlyA 440
Db -----CTCAGGCCCTATGGAG 1173
QY laPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValP 460
Db CCTTCGGTGGCTCCCGCTTTCTGCATCTTCAGTGTCTTTTCACCTTTGTTCTGTGCC 1233
QY roGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db CTGAACATAAGGAAGACTCTGGAAACAATCACAGCCCAATTTTGGGGGGCA 1286

RESULT 3
LOCUS AY414180 1008 bp DNA linear GSS 17-DEC-2003
DEFINITION Homo sapiens SLC2A8 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY414180
VERSION AY414180.1 GI:39770142
KEYWORDS GSS.
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
REFERENCE 1. (bases 1 to 1008) Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2. (bases 1 to 1008) Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
AUTHORS Adams,M.D. and Cargill,M.
TITLE Direct Submision
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering

```

```

FEATURES
    source          them based on alignment.
    Location/Qualifiers
        1..1008
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        <1..>1008
        /gene="SLC2A8"
        /locus_tag="HCW5138"
ORIGIN
Alignment Scores:
Pred. No.:      3,72e-152      Length:      1008
Score:          1713.00        Matches:     332
Percent Similarity: 99.10%      Conservative: 0
Best Local Similarity: 99.10%    Mismatches: 3
Query Match:    69.72%        Indels:     0
DB:             9              Gaps:       0
US-09-886-954A-1 (1-477) x AY414180 (1-1008)
QY 143 ValTyrIleSerGluIleAlaLysProAlaValArgGlyLeuLeuGlySerCysValGln 162
Db 1  GTCATACATCTCCGAAATGCCTACCAGCACTCCGGGGTTCCTCGGCTCTCTGTGTGCG 60
QY 163 LeuMetValValValGlyLeuLeuAlaTyrIleuAlaGlyTrpValLeuGluTrpArg 182
Db 61  CTAATGTCGTCGTCGGCATCTCTCGCTTACCCTGGCAGCGCTGGGTGTGCTGGAGTGG 120
QY 183 TrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPheMet 202
Db 121 TGGTGGCTGTGTGGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTG 180
QY 203 ProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGlnAlaMetAlaAlaLeu 222
Db 181  CCCGAGACCCCGGCTTCTCTGCTGACTCAGCACAGCGCCAGGAGGCCATCGCCCGCTG 240
QY 223 ArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGluGlnSer 242
Db 241  CGGTTCCTGTGGGGCTCCGAGCAGGGCTGGGAAGACCCCGCCATCGGGGGCTGAGCAG 300
QY 243 PheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleLeuGlyValSer 262
Db 301  TTTTCACTGGCCCTGCTGGCGAGCCCGGCTTACAAAGCCCTTCAATCATCGGGGTCTCC 360
QY 263 LeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGluThrIle 282
Db 361  CTGATGGCCCTTCCAGCAGCTGTCTGGGGGTCAAGCCCGCTCATGTCAGAGACCATC 420
QY 283 PheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyValIleGln 302
Db 421  TTTGAAGGCGCAAGTTCAAAGGACAGCAGCCCTGGCTGGGTGGGTGGGTGGGTGGGTG 480
QY 303 ValLeuPheThrAlaValAlaLalaLeuIleMetAspArgAlaGlyArgArgLeuLeuLeu 322
Db 481  GTGTGTTCACAGCTGTGGGGGTCTTCATCATGTGACAGACAGCAGGGCGGAGGCTGCTC 540
QY 323 ValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPheLysLeu 342
Db 541  GTCTTTNNNNGTGTGTCATGTTGTTTCAGCACAGTGCCTTCCGGCGCTACTTTCAGAGCT 600
QY 343 ThrGlnGlyProGlyAsnSerSerHisValAlaLalaSerSerHisValAlaLalaSerAla 362
Db 601  ACCCGGGTGGCCCTGGCAACTCTCGCAGGTGGCCATCTCGGCGGCTGTCTCTGTGACAG 660
QY 363 ProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeuPheIle 382
Db 661  CCTGTTGATGCCAGCCTGGGGCTGGCGTGGCTGGCTGGCGTGGCGAGCATGTGCTTTCATC 720
QY 383 AlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePhePro 402
Db 721  GCCGGCTTTGGGTGGGTGGGGCCCACTCCCTGGCTCTCATGTGCAGAGATCTTCCCT 780

```

Qy 403 LeuHisValIysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPhe 422
 Db 781 CTGCATGTCAAGGGCGTGGCGCAGGCACTGGCTCCTCACCAACTGGCTCATGGCCTTT 840
 Qy 423 LeuValThrIysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTrp 442
 Db 841 CTGCTGACCAAGGAGTTCAGCAGCCTCATGGAGGCTCTCAGGCCCTATGGAGCCTTCTGG 900
 Qy 443 LeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluThr 462
 Db 901 CTTCGCTCCGCTTCTGCACTTCAGTGCCTTTTCACCTTTGTTCTGTGCTCCCTGAACT 960
 Qy 463 LysGlyIysThrLeuGluInIleThrAlaHisPheGluGlyArg 477
 Db 961 AAAGGAAAGACTCTGGAACAATAACACAGGCCATTTTGAGGGGCGA 1005

RESULT 4
 BX395379
 LOCUS BX395379 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
 DEFINITION cDNA clone CS0DD006YPO6 5-PRIME, mRNA sequence.
 ACCESSION BX395379
 VERSION BX395379.2 GI:46875461
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1098)
 AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT On May 13, 2003 this sequence version replaced gi:30620615.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremlieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 5383.f
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?s=CS0DD006DH03QP1&c=5383.f.

FEATURES
 source 1..1098
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DD006YPO6"
 /tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
 /notes="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and EcoR V sites of the pCMVSPORT 6 vector.
 Library was normalized."/>
 ORIGIN
 Alignment Scores:
 Pred. No.: 6,21e-131 Length: 1098
 Score: 1490.50 Matches: 320
 Percent Similarity: 68.72% Conservative: 3
 Best Local Similarity: 68.09% Mismatches: 23
 Query Match: 60.66% Indels: 127
 DB: 5 Gaps: 1

US-09-886-954A-1 (1-477) x BX395379 (1-1098)
 Qy 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProGlyGlySerAla 20
 Db 19 ATGACGCCGAGGACCCAGAGGAAACCCAGCCGCTTCTGGGGCCTCTCGCGCGCAG- 74

Qy 21 ProArgGlyArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
 Db 74 ----- 74
 Qy 41 GlyPheAlaLeuGlyTyrSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 Db 74 ----- 74
 Qy 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
 Db 74 ----- 74
 Qy 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
 Db 74 ----- 74
 Qy 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 Db 74 ----- 74
 Qy 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 Db 74 ----- 74
 Qy 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValAlaArgGlyLeuLeuGlySerCys 160
 Db 75 -----GTCTACATCTCCGAAATCGCTACCCAGCAGTTCGGGGGTTGCTCGGCTCTGT 128
 Qy 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
 Db 129 GTGCAGCTAATGGTCGTGGCAFCCTATG-GCCTACTGGCAGGCTGGKTHRTGGAG 187
 Qy 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 Db 188 TKGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 247
 Qy 201 PheMetProGluThrProArgPheLeuLeuLeuThrGlnHisArgGlnGluAlaMetAla 220
 Db 248 TTTATGCCCGAGACCCCGGCTTCTGCTGACTCAGCAGCAGCCCGCAGGAGGCAATGGCC 307
 Qy 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProIleGlyAlaGlu 240
 Db 308 GCCTTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 367
 Qy 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProCgIleTyrLysProPheIleIleGly 260
 Db 368 CAGAGCTTTCACCTGGCCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 427
 Qy 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
 Db 428 GTCTCCCTGATGGCTTCCAGCAGGCTTTGGGGGTTAACGCCGTTATGTTCTATGCGAG 487
 Qy 281 ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVal 300
 Db 488 ACCATCTTTGAAGAGCCCAAGTTCAGGACAGCAGCCTGGCTGGCTGGCTGGCTGGCTGG 547
 Qy 301 IleGlnValLeuPheThrAlaValAlaAlaLeuLeuMetAspArgAlaGlyArgArgLeu 320
 Db 548 ATCCAGGTCTGTWCACAGCTGTAGCGGCTCYVATCATGACAGCAGCGCCGAGGCTG 607
 Qy 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
 Db 608 CTCCCTGGCTTGTGAGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 667
 Qy 341 LysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer 360
 Db 668 AAGCTGACCCAGGGTGGCCCTGGCAACTCCTCGCAGTGGCCATMTMGCGCGCTGTTC 727
 Qy 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu 380
 Db 728 GCACAGCCTTGTGATGCCAGCTGGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 787
 Qy 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIle 400

Db TWATCGCGCGCTTTGGGCTGGGCTGGGGCCCATCCCGTGGCTCCTCATGTHRGAANTC 847
 QY PheProLeuHisValLysGlyValAlaThrGlyLeCysValLeuThrAsnTrpLeuMet 420
 Db TTCCTCTGCATGTTAAAGGGGCTKGCACAGGATCTCGCTTCTCACCACCTGGCTCATG 907
 QY 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
 Db GCCTTTCTCGTACCAAGGAGTTCAGCAGCTTCATGGAGGTCCTCAGGCCCTATGGAGCC 967
 QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
 Db TTCTKGTCTCCCTCGCTTCTGCATCTCCAGTTT-CTTTTCATTGTTCTTGTCTCT 1026
 QY 461 -GluThrLysGlyLysThrLeuGluIn 469
 Db TAAACTAAAGGAAATTTTGGACCAA 1054

AY414182 1000 bp DNA linear GSS 17-DEC-2003
 Mus musculus SLC2A8 gene, VIRTUAL TRANSCRIPT, Partial sequence,
 Genomic survey sequence.
 AY414182
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
 Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,
 Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Sniinsky, J.J.,
 Adams, M.D. and Cargill, M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 Science 302 (5652), 1960-1963 (2003)
 14671302
 2 (bases 1 to 1000)
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
 Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,
 Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Sniinsky, J.J.,
 Adams, M.D. and Cargill, M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment
 Location/Qualifiers
 1. .1000 Length: 1000
 /organism="Mus musculus" Matches: 284
 /mol_type="genomic DNA" Conservative: 22
 /db_xref="taxon:10090" Mismatches: 26
 <1..>1000 Indels: 4
 /gene="SLC2A8" Gaps: 2
 /locus_tag="HCM5138"

Alignment Scores:
 Pred. No.: 1-62e-130 Length: 1000
 Score: 1485.50 Matches: 284
 Percent Similarity: 91.07% Conservative: 22
 Best Local Similarity: 84.52% Mismatches: 26
 Query Match: 60.46% Indels: 4
 DB: 9 Gaps: 2

US-09-886-954A-1 (1-477) x AY414182 (1-1000)

QY 143 ValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGln 162
 Db 1 GTTTACATCTCGGAAATCGCTACCAGCTCCGAGGACTGCTCGGCTCCTGTGTGCAG 60

QY 163 LeuMetValValIleGlyLeuLeuAlaTyrLeuAlaGlyTrpValLeuGluTrpArg 182
 Db CTGATGTTGTCACTGGCATCTCTGGCTATATGGCAGGCTGGTCTTAGAGTGGCG 120
 QY 183 TrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPheMet 202
 Db TGGTGGCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 180
 QY 203 ProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaAla 222
 Db CCCGAGACCCACGTTTCTCTCACTCAACACAGTACCAGAGGCCATGGCTGCCCTG 240
 QY 223 ArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProIleGlyValAlaGluInSer 242
 Db CGCTTCTCTGGGCTCTGAGGAGGCTGGGAAGACCCCTGTTGGGGCTGAGCACGGC 300
 QY 243 PheHisLeuAlaLeuArgGlnProGlyIleTyrLysProPheIleIleGlyValSer 262
 Db TTCCAGCTGGCCCTCTGAGGCGCCCTGGCATCTACAGCCCTCATCATCGGCATTTCC 360
 QY 263 LeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGluThrIle 282
 Db CTCATGGTCTTCCAGCAGCTGTCAAGGCTCACTGATCATGTTCTATGCCAACAGCATC 420
 QY 283 PheGluGluAlaLysPheLysAspSerSer-LeuAlaSerValValValGlyValIleG 302
 Db TTCGAGGAGGCCAAGTTCAAG-----CAGCCTGGCCTCGGTCACTGTGGCATATAATCA 474
 QY 302 nValLeuPheThrAlaValAlaLeuLeuMetAspArgAlaGlyArgArgLeuLeuLe 322
 Db GGTCTGTCACTGCTGTGGCGCCCTCATCATGGACAGACAGCGGGGAGGCTGCTCCT 534
 QY 322 uValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPheLysLe 342
 Db GGCCTTGTGGGCTGATCATGTGTTCAGTATGAGTGCCTTTGGTACCTACTTCAACT 594
 QY 342 uThrGlnGlyGlyProGlyAsnSerHisValAlaIleSerAlaProValSerAlaG 362
 Db GACCCAGAGCCTCCAGCAACTCTCCACAGTAGGCGCTG---GTGCCCATCGCGGGGA 651
 QY 362 nProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeuPheI 382
 Db GCCTGTGGATGTCACAGTGGGACTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 711
 QY 382 eAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePhePr 402
 Db TGCTGGCTTTGGGCTGGGCTGGGCACTCCCTGGCTCCTCATGTTCAGAGATCTTCCC 771
 QY 402 oLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPh 422
 Db TCTGCAATGTCAGGGTGGCTACCGGCATCTGTGTCCTCCACCACTGGTTCAATGGCCT 831
 QY 422 eLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTr 442
 Db TCTAGTGCACCAAGAGTTTCAGCAGCGTCAATGGAGATGTCTCAGACCTTACGGTCTTG 891
 QY 442 pLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluTh 462
 Db GCTCACCGCTGGCTTCTGGCTCCTCAGTGTCTTATTCACACTGACCGTGTCTCCCTGAGC 951
 QY 462 rLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 Db TAAAGGCAGGACTCTGGACCAAGTACAGCCCACTTCGAGGGACGA 997

RESULT 6
 EM545247
 LOCUS EM545247 1085 bp mRNA linear EST 20-FEB-2002
 DEFINITION AGENCOURT 6497266 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5726945
 5', mRNA sequence.
 ACCSSION EM545247
 VERSION EM545247
 KEYWORDS EM545247.1 GI:18777177
 EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1085)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Invitrogen
 cDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Plate: LLM12720 row: d column: 18
 High quality sequence stop: 679.
 Location/Qualifiers
 1..1085
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5726945"
 /tissue_type="hippocampus"
 /lab_host="DH10B"
 /clone_lib="NIH MGC 124"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: EcoRV (destroyed); Site 2: NotI; RNA source male hippocampus, age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 0.9-4 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 012."
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,47e-120 Length: 1085
 Score: 1379.00 Matches: 279
 Percent Similarity: 87.77% Conservative: 1
 Best Local Similarity: 87.46% Mismatches: 6
 Query Match: 56.13% Indels: 33
 DB: 4 Gaps: 3
 US-09-886-954A-1 (1-477) x BMS45247 (1-1085)
 QY 162 GlnLeuMetValValGlyLeuLeuLeuAlaTyrLeuAlaGlyTrpValLeuGluTrp 181
 Db 1 CAGCTAATGGTCGCTGGCATCTCTCTGGCTACCTGGCAGGCTGGAGTGG 60
 QY 182 ArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPhe 201
 Db 61 CGCTGGCTGGCTGTGTGGGCTGCGTGGCCCTCCCTCATGCTCTCATGTGCTTC 120
 QY 202 MetProGluThrProArgPheLeuLeuThrGlnHisArgGlnGluAlaMetAlaAla 221
 Db 121 ATGCCAGACCCCGCGCTTCTGCTGACTAGCAGCGGCGCAGAGCCGAGCCGCC 180
 QY 222 LeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProGlyAlaGluGln 241
 Db 181 CTGCGGTTCTGTGGGCTCCGAGCAGGGCTGGGAAGACCCCTCCATCGGGCTGAGCAG 240
 QY 242 SerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleGlyVal 261
 Db 241 AGCTTTACCTGGCCCTGTGGGCGCAGCCCGGCATCTACAAGCCCTTCATCATCGGCGTC 300
 QY 262 SerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGluThr 281
 Db 301 TCCCTGATGGCTTCCAGCAGCTGTCCGGGGTCAACCGCTCATGTTCATCAGAGACC 360
 QY 282 IlePheGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyValIle 301

Db 361 ATCTTTGAGAGGCCCAAGTTCAAGGACAGCAGCCTGGCTCGTCTGGGTGTCAATC 420
 QY 302 GlnValLeuPheThrAlaValAlaLeuLeuMetAspArgAlaGlyArgLeuLeu 321
 Db 421 CAGGTGCTGTTTCAAGCTGTGGCGCTCTCATATGGACAGAGCAGGGGGAGGCTGCTC 480
 QY 322 LeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPheLys 341
 Db 481 CTGGTCTTGTCAAGTGTGGTGTGTTTTCAGCAGGAGTGCCTTCGGCGCTACTTCAAG 540
 QY 342 LeuThrGlnGlyProGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSerAla 361
 Db 541 CTGACCCAGGGTGGCCCTGGCAACTCTCGCACGTGGCCATCTCGGGCGCTGTCTCTGCA 600
 QY 362 GlnProValAspAlaSerValGlyLeuAlaTyrLeuAlaValGlySerMetCysLeuPhe 381
 Db 601 CAGCCTGTGTATGCCAGCGTGGGGCTGGCTGGCTGGTGGCGAGCAATGTGCTCTTC 660
 QY 382 IleAlaGlyPheAlaVal-GlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePhe 401
 Db 661 ATCCGCGGCTTTGGGCTGGGCTGGGGCCCATCCCTGGCTCTCATGTCTCAGAGATCTT 720
 QY 401 eProLeuHisValLysGlyValAlaThrGlyIleCysValLeuLeuThrAsnTrpLeuMetAl 421
 Db 721 CCTCTGCAATGTCAAGGGCGTGGCGACAGGCACTCTCGCTCTCCACCAACTGGCTCATGGC 780
 QY 421 aPheLeuValThrLysGlyPheSerSerLeuMet----- 432
 Db 781 CTTTCTCTGACCAAGGAGTTTCAGCAGCCTCAATGCTCTGCTGCTCTGAGGACTCAN 840
 QY 433 -----GluValLeuArgPro----- 437
 Db 841 GAACACCTTCAGACTTTCAGACCTTCGGGTTCAGCCCTCCATGGCGCAAGATAAAGCAACGG 900
 QY 438 -----TyrGly-----AlaPheTrpLeuAlaSerAlaPhe 447
 Db 901 NAAAAAGGAGGGGGGCTCCAAAGGACCTTTGGCTTTTGGCTTGAAGGGCCCTTT 955
 RESULT 7
 BG479842
 LOCUS 60252737F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:4650906 5',
 DEFINITION mRNA sequence.
 ACCESSION BG479842
 VERSION BG479842.1 GI:13412121
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 932)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Plate: LLM1433 row: a column: 19
 High quality sequence stop: 820.
 Location/Qualifiers
 1..932
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4650906"
 /tissue_type="choriocarcinoma"
 FEATURES
 source

/lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_21"
 /note="Organ: Placenta; Vector: pORB7; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGACAGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Alignment Scores:	1.02e-116	Length:	932
Pred. No.:	1340.00	Matches:	293
Score:	93.93%	Conservative:	1
Percent Similarity:	93.61%	Mismatches:	17
Best Local Similarity:	54.54%	Indels:	7
Query Match:	4	Gaps:	0

US-09-886-954A-1 (1-477) x BG479842 (1-932)

QY	51	IleProSerLeuGlnArgAlaAlaProProAlaProArgLeuAspAlaAlaSer	70
Db	2	ATCCCTAGCCTGCAGCGCCCGCGCCCGCGCCCGCGCGCTGGACGACGCGCCGCTCC	61
QY	71	TrpPheGlyAlaValThrLeuGlyAlaAlaGlyValLeuGlyGlyTrpLeu	90
Db	62	TGGTTCCGGGCTGTCTGACCCCTGGTGCCTGGCGGGGGAGTGTCTGGCGGCTGGCTG	121
QY	91	ValAspArgAlaGlyArgLysLeuSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	110
Db	122	GTGGACCGCGCGGGCGCAAGCTGACCCCTTTGCTGTGCTGCTGCTGCTGCTGCTGCTG	181
QY	111	PheAlaValIleThrAlaAlaGlnAspValTrpMetLeuLeuGlyValArgLeuLeuThr	130
Db	182	TTTTCCGCTCATCCCGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	241
QY	131	GlyLeuAlaCysGlyValAlaSerLeuValAlaProValTrpIleSerGluIleAlaTrp	150
Db	242	GGCTGGCGCTGGCGGTGTCCCTCCCTAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	301
QY	151	ProAlaValArgGlyLeuLeuGlySerCysValGlnLeuMetValValGlyIleLeu	170
Db	302	CCAGCAGTCCGGGGTGTCTCGGCTCCTGTGTGCAGCTAATGGTGTGTGTGTGTGTGTGTGT	361
QY	171	LeuAlaTrpLeuAlaGlyTrpValLeuGluTrpArgTrpLeuAlaValLeuGlyCysVal	190
Db	362	CTGGCCTACCTGGCAGCTGGGTGTCTGGAGTGGCGCTGGCTGGCTGGCTGGCTGGCTGGCT	421
QY	191	ProProSerLeuMetLeuLeuLeuMetCysPheMetProGluThrProArgPheLeuLeu	210
Db	422	CCCCCTCCCTCATGCTGCTTCTCATGCTTCTCATGCTCCGCGAGACCCCGCGCTCCCTGCTG	481
QY	211	ThrGlnHisArgArgGlnGluAlaMetAlaAlaLeuArgPheLeuTrpGlySer-GluG1	230
Db	482	ACTCAGCACAGCGCCCGAGGCGCCATGGCCCGCTGCGGTTCTGTGGGGCTCCAGACA	541
QY	230	nGlyTrpGluAspProProIleGlyAlaGluGlnSerPheHisLeuAlaLeuLeuArgG1	250
Db	542	GGGCTGGGAAGACCCCGCCCG	601
QY	250	nProGlyIleTrpLysProPheIleIleGlyValSerLeuMetAlaPheGlnGlnLeuSe	270
Db	602	GCCCGGCATCTCAAGCCCTTCATCATCGGCGCTCCCTGATGGCGCTCCACGACGCTGTC	661
QY	270	rGlyValAsnAlaValMetPheTrpAlaGluThrIlePheGluAlaLysPheLysAs	290
Db	662	GGGGGTCAAGCGCGCTCTGTTCATGTCAGAGACCATCTTTGAGAGGCGCCAGTTCAAGGA	721
QY	290	pSerSerLeuAla-SerValValValGlyValIleGlnValLeuPheThrAlaValAlaA	310
Db	722	CAGCAGCCTGGCGCTCGT	781

QY	310	laLeuIleMetAspArgAlaGlyArgLeuLeuLeuValLeuSerGlyValValMetV	330
Db	782	CTCTCATCATGACAGAGCAGCGCGG-AGGCTGGTCTGCTGGCTGG-TCAGGTGGCCANGG	839
QY	330	alPheSerThrSerAlaPheGlyAlaTrpPheLysLeuThrGlnGlyGlyProGlyAsnS	350
Db	840	TGTGAGCGCGAGTGCCTC-GGCGCTTAATTCAGCTGACCCG-GGTTGGCTGGCAATT	897
QY	350	erSerHisValAlaIleSerAlaProValSerAla	361
Db	898	CCTCGCC-GTGGCCATCTCGGGCCTGTCTCTGGC	931

RESULT 8
 BG700749
 LOCUS 602681616F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4814451 5',
 DEFINITION mRNA sequence.
 ACCESSION BG700749.1 GI:13970402
 VERSION BG700749
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 830)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: rsb@nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LRAM10709 row: p column: 04
 High quality sequence stop: 789.
 Location/Qualifiers
 source
 1.830
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4814451"
 /tissue_type="hippocampus"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_95"
 /note="Organ: brain; Vector: pBluescriptR (modified
 pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
 (gtcgag); Oligo-dT primed using primer
 5'-TTTTTTTTTTTTTTVN-3', size-selected for average
 insert size 2.5 kb and normalized to ROT 5. This is a
 primary library enriched for full-length clones and
 constructed using the Cap-trapper method (Carninci, in
 preparation). Library constructed by M. Brownstein
 (NIH/NHGRI, National Institutes of Health). Note: this
 is a NIH_MGC Library."

ORIGIN

Alignment Scores:	2.54e-112	Length:	830
Pred. No.:	1293.00	Matches:	265
Score:	96.01%	Conservative:	0
Percent Similarity:	96.01%	Mismatches:	8
Best Local Similarity:	52.63%	Indels:	3
Query Match:	4	Gaps:	0

US-09-886-954A-1 (1-477) x BG700749 (1-830)

QY	133	AlaCysGlyValAlaSerLeuValAlaProValTrpIleSerGluIleAlaTrpProAla	152
----	-----	--	-----

Db 5 GCCTGGGGTGTGCTCCCTAGTGGCCCGGCTACATCTCCGAAATCGCCTACCCAGCA 64
 QY 153 ValArgGlyLeuLeuGlySerCysValGlnLeuMetValValGlnLeuLeuLeuA 172
 Db 65 GTCCGGGGTGTGCTGGCTCCTGCTGTGCGACTAATGCTGTGCTGGCATCTCTCCGGCC 124
 QY 173 TyrLeuAlaGlyTrpValLeuGluTrpAraGfrTrpLeuAlaValLeuGlyCysValPro 192
 Db 125 TACCTGGCAGCTGGGTGTGAGTGGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 184
 QY 193 SerLeuMetLeuLeuMetCysPheMetProGluThrProArgPheLeuLeuThrGln 212
 Db 185 TCCTCATGCTGCTTCTCATGTGCTTTCATGCCCGAGACCCCGGCTTCTGTGACTCAG 244
 QY 213 HisArgGlnGlnAlaMetAlaLeuAraGlnPheLeuTrpGlySerGluGlnGlyTrp 232
 Db 245 CACAGCGCCGAGGAGCCATGGCCGCTTCTGTGGGCTCCGAGCAGGGCTGG 304
 QY 233 GluAspProIleGlyAlaGluGlnSerPheHisLeuAlaLeuLeuArgGlnProGly 252
 Db 305 GAAGACCCCCATCGGGCTGAGCAGAGCTTTCACCTGGCCCTGTGGCCAGCCCGC 364
 QY 253 IleTyrLysProPheIleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyVal 272
 Db 365 ATCTACAAGCCCTTCATATPCGGCGCTTCCCTGATGGCTTCCAGCAGCTGTGGGGGTC 424
 QY 273 AsnAlaValMetPheTyrAlaGluThrIlePheGluGluAlaLysPheLysAspSerSer 292
 Db 425 AACCCGCTCATGTTCTATGACGAGACCACTTTTGAAGAGCCCAAGTTCAAGACAGCAGC 484
 QY 293 LeuAlaSerValValGlyValIleGlnValLeuPheThrAlaValAlaLeuIle 312
 Db 485 CTGGCTCGCTGCTGGTGTTCATCCAGGTGCTTTCACAGCTGTGGCGCTCTCATC 544
 QY 313 MetAspArgAlaGlyArgLeuLeuLeuValLeuSerGlyValValMetValPheSer 332
 Db 545 ATGGACAGACGAGCGGGAGGCTGCTTCTGGTCTTGTAGGTGTGCTGTGGTTCAGC 604
 QY 333 ThrSerAlaPheGlyValTyrPheLysLeuThrGlnGlyProGlyVAsnSerSerHis 352
 Db 605 ACAGTGGCTTGGCCCTTCTCAAGCTGACCCAGGGTGGCCCTGGCACTCTCCGCAC 664
 QY 353 ValAlaIleSerAlaProValSerAla-GlnProVal-AspAlaSerValGlyLeuAlaT 372
 Db 665 GTGGCCATCTGGCGCTGCTCTGCCACAGCCTGTTTGTATGCCAGCGTGGGCTGGCT 724
 QY 372 rLeu-AlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyPro 391
 Db 725 GGCTGGGCGCGTGGCAGCATGCTCTTTCATCGCGCTTGGGCTGGGCTGGGCGCC 784
 QY 392 IleProTrpLeuLeuMetSerGluIlePheProLeuHisVal 405
 Db 785 ATCCCTTCTCTCATGTCAGAGATTTTCCCTCTGTGCTGTC 826

RESULT 9
 BI334832
 LOCUS 991 bp mRNA linear EST 30-JUL-2001
 DEFINITION 60298959F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5141007 5', mRNA sequence.
 ACCESSION BI334832
 VERSION BI334832.1 GI:15019489
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 991)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E.B. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM11347 row: f column: 16
 High quality sequence stop: 835.
 Location/Qualifiers
 1..991
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5141007"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_12"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.4 kb. Library prepared by Life Technologies."
 Alignment Scores:
 Pred. No.: 5.66e-110 Length: 991
 Score: 1269.50 Matches: 290
 Percent Similarity: 65.05% Conservative: 6
 Best Local Similarity: 63.74% Mismatches: 21
 Query Match: 51.67% Indels: 140
 DB: 4 Gaps: 6
 US-09-886-954A-1 (1-477) x BI334832 (1-991)

ORIGIN
 QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
 Db 17 ATGACGCCCGAGACCCAGAGAAACCCAGCCGCTTCTGGGGCTCTGGCGGCAG---- 72
 QY 21 ProAspGlyArgArgValPheLeuAlaAlaPheAlaAlaAlaLeuGlyProLeuSerPhe 40
 Db 72 ----- 72
 QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 Db 72 ----- 72
 QY 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
 Db 72 ----- 72
 QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
 Db 72 ----- 72
 QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 Db 72 ----- 72
 QY 121 TrpMetLeuLeuGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 Db 72 ----- 72
 QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
 Db 73 -----GCTACATCTCCGAAATCGCCTACCCAGCAGTCCGGGGTGTGCTGGCTCTGT 126
 QY 161 ValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
 Db 127 GTGCAGCTAATGCTGTGCTGGCATCTCTGGGCTACCTGGGCTGGGCTGGGCTGGGAG 186
 QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 Db 187 TGGCGCTGGCTGTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 246

201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
 Db 247 TTCATGCCCCAGACCCGGCTTCCTGCTACTCAGCACAGCGCCGAGAGCCATGCGC 306
 Qy 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProIleGlyValaGlu 240
 Db 307 GCCTTGGGTTCTGTGGGCTCCGAGCAGGGCTGGGAAGACCCCCCATCGGGCTGAG 366
 Qy 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIleGly 260
 Db 367 CAGAGCTTTCACCTGCCCTGTGGCCAGCCCGGCTCATCAAGCCCTTCATCATCGGC 426
 Qy 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
 Db 427 GTCTCCCTGATGGCTTCCAGCAGCTGTCCGGGGTCAACCGCGTCAATGTTCTATGCAGG 486
 Qy 281 ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVal 300
 Db 487 ACCAICTTTCAGAGGCCAAGTTCAGGACAGCAGCCCTGGCTCGTGGTGGTGTGTC 546
 Qy 301 IleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArgArgLeu 320
 Db 547 ATCCAGGCTGCTTCACAGCTGTGGCGCTCTCATCATGACAGCAGCGGGAGGCTG 606
 Qy 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
 Db 607 CTCTCTGGTCTGTGTCAGGTGTGGTCAATGGTTCAGCAGAGTGCCTTCGGCCCTACCTC 666
 Qy 341 LysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProValIse 360
 Db 667 AGCTGACCCAGGGTGGCTTGGCAACTCTCGCAACGTGGCCATCTGGGCGCTGTCTC 726
 Qy 360 rAlaGln-ProValAspAlaSerVal-GlyLeuAlaTrpLeuAlaVal-GlySerMetCy 379
 Db 727 TGCAACAAGCTGTGTATGTCAGCGTGGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 786
 Qy 379 sLeuPheIleAlaGlyPheAlaValGly--TrpGlyProIleProTrpLeu---LeuMe 397
 Db 787 CTCCTTCATCGCCGGCTTGGGTTGGGCTTGGGGGCCATTCCTCCCTGGTCCCTCATTTG 846
 Qy 397 tSerGluIlePheProLeuHisValLysGlyValAlaThrGly-----II 412
 Db 847 TCAGAAGACTTTCCTCCCTC-----TGCAATTGTCAGGGCGCTGGCCACAGGGCATT 897
 Qy 412 eCysValLeuThrAsnTrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMe 432
 Db 898 CTAGGTCCTCCACAATGGTCCATGGCTTCCCGGGACCAG-GAGTTCAGCCAGCTCTT 956
 Qy 432 tGluValLeuArgProTyrGlyAlaPheTrpLeuAla 444
 Db 957 GGAAGTCCCGCCCTTGAC---CTCTGGGTGTCT 990
 RESULT 10
 AY414181 1004 bp DNA linear GSS 17-DEC-2003
 LOCUS Pan troglodytes SLC2A8 gene, VIRTUAL TRANSCRIPT, partial sequence,
 DEFINITION genomic survey sequence.
 ACCESSION AY414181
 VERSION AY414181.1 GI:39770143
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM
 Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
 REFERENCE 1 (bases 1 to 1004)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
 Todd,M.A., Tanenbaum,D.M., Civeello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sainsky,J.J.,
 Adams,M.D. and Cargill,M.

14671302
 2 (bases 1 to 1004)
 REFERENCE Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
 AUTHORS Todd,M.A., Tanenbaum,D.M., Civeello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sainsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 Location/Qualifiers
 source 1..1004
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 <1..>:1004
 /gene="SLC2A8"
 /locus_tag="HCMS138"
 ORIGIN
 Alignment Scores: Length: 1004
 Pred. NO.: 6.43e-110 Matches: 263
 Score: 1269.00 Conservative: 3
 Percent Similarity: 79.88% Mismatches: 67
 Best Local Similarity: 78.98% Indels: 0
 Query Match: 51.65% Gaps: 0
 DB: 9
 US-09-886-954A-1 (1-477) x AY414181 (1-1004)
 Qy 145 IleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGlnLeuMet-164
 Db 3 ATCTCCGAATCGCTACCCAGCAGTCCGGGGTTGCTCGGCTCTTGTGTGCAGCTGATG 62
 Qy 165 ValValValGlyIleLeuLeuAlaTyrIleAlaGlyTrpValLeuGluTrpArgTrpLeu 184
 Db 63 GTCGTCGTGGGCATCTCTGGCCCTACCTGGCAGNNNNNNNNNNNNNNNNNNNNNNN 122
 Qy 185 AlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPheMetProGlu 204
 Db 123 NNN 182
 Qy 205 ThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaLeuArgPhe 224
 Db 183 NNN 242
 Qy 225 LeuTrpGlySerGluGlnGlyTrpGluAspProIleGlyAlaGluGlnSerPheHis 244
 Db 243 NNN 302
 Qy 245 LeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIleGlyValSerLeuMet 264
 Db 303 CTGGCCCTGCTGGCGAGCCCGGCATCTACAGGCCCTTCGTCATCGGGCTCCCTCGATG 362
 Qy 265 AlaPheGlnLeuLeuSerGlyValAsnAlaValMetPheTyrAlaGluThrIlePheGlu 284
 Db 363 GCCTTCCAGCAGCTGTGGGGGTCAACCGCTCATGTTCTATGCAGAGCACCATCTTTGAA 422
 Qy 285 GluAlaLysPheLysAspSerSerLeuAlaSerValValGlyValIleGlnValLeu 304
 Db 423 GAGGCCAAGTTCAGGACAGCAGCGCTGGCTCGGTCFCGCGGTGTCTCCAGTGTCTG 482
 Qy 305 PheThrAlaValAlaLeuIleMetAspArgAlaGlyArgArgLeuLeuLeuValLeu 324
 Db 483 TTCACAGCTGTGGGCTCTCATCATGACAGCAGCGGGGAGGCTGCTCTCTGTCTG 542
 Qy 325 SerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPheLysLeuThrGln 344
 Db 543 TCAGGTGTGGTCAATGGTTCAGCAGCAGTGCCTTCGGCGCTACTTCAAGCTGACCCAG 602
 Qy 345 GlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSerAlaGlnProVal 364

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 880)
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Cloned through the I.M.A.G.E. Consortium/LLNL at:
Plate: LLCM1930 row: k column: 22
High quality sequence stop: 869.
Location/Qualifiers
1. 880
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5447181"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pORB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

Alignment Scores:
Pred. No.: 2,19e-107 Length: 880
Score: 1241.50 Matches: 280
Percent Similarity: 58.82% Conservative: 0
Best Local Similarity: 58.82% Mismatches: 5
Query Match: 50.53% Indels: 193
DB: 4 Gaps: 2

US-09-886-954A-1 (1-477) x BM044230 (1-880)

QY 4 GluAspProGluThrGlnProLeuLeuGlyProGlySerAlaProArgGly 23
DB 4 GAGGACCCAGAGAAACCCAGCGCTTCGGGGCTTCCTGGGGCAG-----50
QY 24 ArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPheAla 43
DB 50 -----50
QY 44 LeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaProArg 63
DB 50 -----50
QY 64 LeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaAlaGly 83
DB 50 -----50
QY 84 GlyValLeuGlyTyrTrpLeuValAspArgAlaGlyArgGlyLeuSerLeuLeuLeuCys 103
DB 50 -----50
QY 104 SerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMetLeu 123
DB 50 -----50
QY 124 LeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaProVal 143

Db 51 -----GTC 53
QY 144 TyrIleSerGluLeuAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGlnLeu 163
Db 54 TACATCTCCGAAATCGCCTACCCAGCAGATCCGGGGTTGCTGGCTCTCTGTGTGACGTA 113
QY 164 MetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGluTrpArgTrp 183
Db 114 ATGTGCTGTGCGACATCCCTCGGCTTACTGGCA-----149
QY 184 LeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPheMetPro 203
Db 149 -----149
QY 204 GluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaAlaLeuArg 223
Db 149 -----149
QY 224 PheLeuTrpGlySerGluGlnGlyTrpGluAspProIleGlyAlaGluGlnSer-Ph 243
Db 150 -----GAGCTT 155
QY 243 eHisLeuAlaLeuLeuArgGlnProGlyIleTyrIleLysProPheIleLeuGlyValSerLe 263
Db 156 TCACCTGGCCCTGTGGCGCAGCCCGGCATCTACAGCCCTTCATCATCGGCTCTCCCT 215
QY 263 uMetAlaPheGlnGlnLeuSerGlyValAenAlaValMetPheTyrAlaGluThrIlePh 283
Db 216 GATGGCTTCAGCAGCTGTGGGGTCAACCGCTCATGTTCTATGACAGACCACTT 275
QY 283 eGlnGluAlaLysPheLysAspSerLeuAlaSerValValValGlyValIleGlnVa 303
Db 276 TGAAGAGCCCAAGTTCAAGGACAGCAGCCCTGGCTCGTCTGGGTGTCACTCCAGGT 335
QY 303 lLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeuLeuLeuVa 323
Db 336 GCTCTTCACAGCTGTGGGGCTCTCATCATGGACAGACAGGGCGGGTGTCTCTGGT 395
QY 323 lLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPheLysLeuTh 343
Db 396 CTTGTCAAGTGTGTGTCAAGGTTTCAGCAGCAGTCCCTTCGGCCCTTCTCAAGCTGAC 455
QY 343 rGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSerAlaGlnPr 363
Db 456 CCAGGCTGGCCCTGGCAACTCCTGCACGTGGCCATCTCGGGCTGTCTGTGCACAGCC 515
QY 363 oValAspAlaSerValGlyLeuAlaTrpLeuAlaValIleLysSerMetCysIleuPheIleAl 383
Db 516 TGTGTATGCCAGCGTGGGGCTGGCTGGCTGGCGGCGAGCATGCGCTTTCATCGC 575
QY 383 aGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePheProLe 403
Db 576 CGCTTTTGGGTGGCTGGGGCCCATCCCTGGCTCTCATGTTCAGAGATCTTCCCTCT 635
QY 403 uHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPheLe 423
Db 636 GCATGTCAAGGGCTGGCCAGCAGGCACTGGCTTCACCACTGGCTTCATGGCTTCT 695
QY 423 uValThrIleGluPheSerLeuMetGluValLeuArgProTyrGlyAlaPheTrpLe 443
Db 696 CGTGACCAAGAGATTTCAGCAGCCCTCATGGAGGTCTTCAGGGCCCTATGAGGCTTCTGGCT 755
QY 443 uAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCys-ValProGluThrL 463
Db 756 TGCCCTCCGTTCTGCATCTTCAGTGTCTTTTTCATTTTGGTGTCTCCGAACTA 815
QY 463 ysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db 816 AAGGAAAGACTCTGGAACAAAATCACAGCCA--TTTGGGGGGCGA 857

RESULT 13
BG478000

LOCUS BG478000 828 bp mRNA linear EST 21-MAR-2001
 DEFINITION 602522606F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4641145 5',
 mRNA sequence.

ACCESSION BG478000
 VERSION BG478000.1 GI:13410279
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>;
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC/DCTD/bnp
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM1407 row: k column: 02
 High quality sequence stop: 788.

FEATURES
 source Location/Qualifiers
 1..828
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4641145"
 /issue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_20"
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Alignment Scores:
 Pred. No.: 4, 32e-107 Length: 828
 Score: 1238.00 Matches: 250
 Percent Similarity: 95.08% Conservative: 1
 Best Local Similarity: 94.70% Mismatches: 8
 Query Match: 50.39% Indels: 5
 DB: 4 Gaps: 0

US-09-886-954A-1 (1-477) x BG478000 (1-828)

QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProGlyGlySerAla 20
 |||||
 Db 47 ATGACGCCGAGACCCAGAGAAACCCAGCGCTTCTGGGGCTCTGGCGCGCGG 106
 QY 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
 |||||
 Db 107 CCCCAGCGCGCGCGCTTCTCGCGCGCTTGGCGCTGGCCGCGCCACTCAGCTTC 166
 QY 41 GlyPheAlaLeuGlyTyrSerProAlaIleProSerLeuGlnArgAlaProPro 60
 |||||
 Db 167 GCGTTCGCGCTGGCTACAGCTCCCGCGCAICCCCTAGCTGACGCGCGCGCGCC 226
 QY 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValThrLeuGlyAla 80
 |||||
 Db 227 GCCCGCGCTGCACGACGCGCGCGCTCTCGTGGGGCTCTCGTACCCCTGGGTGCC 286
 QY 81 AlaAlaGlyValLeuGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
 |||||
 Db 287 GCGCGCGGGGAGTGTGGCGCGCTGGCTGGTGGACCGCGCGCGCGCGCAAGCTGAGCCTC 346

QY 101 LeuLeuCySerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 |||||
 Db 347 TTGCTGTGCTCCGTCGCTTCCGCGGCTTGGCGTTCATCACCGCGGCCAGGACGTG 406
 QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 |||||
 Db 407 TGGAGTGTGCTGGGGGGCGGCTCCTCACCGGCTGGCGCTGGCGTGTGGCTCCCTAGTG 466
 QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
 |||||
 Db 467 GCCCGGCTTACATCTCCGAAATCGCTTACCCAGCAGTCCGGGGGTGTCTGGCTCCTGT 526
 QY 161 ValGlnLeuMetValValGlyIleLeuLeuAla--TyrLeuAlaGlyTrpValLeuG 180
 |||||
 Db 527 GFGAGCTAATGTCGTGTCGCGCATCTCTTGGCCCTTACCTGGCAGGCTGGGTCTGG 586
 QY 180 LuTrpArgTrpLeuAlaValLeuGlyCysVal-ProProSerLeuMetLeuLeuMet 199
 |||||
 Db 587 AGTGGCGCTGGCTGGCTGTGCTGGGCTGGCTGGCCCTTCCCTCATGTGCTTCTCATG 646
 QY 200 CysPheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMet 219
 |||||
 Db 647 TGCCTTATGCCGAGACCCGAGCTTCTGCTGACTCACACAGGGCCGAGGAGCCATG 706
 QY 220 AlaAlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAla 239
 |||||
 Db 707 GCGCGCTGGGCTTCTGTGAGGCTCCGAGCAGGCTGGGAGAGACCCCCCATCGGGCT 766
 QY 240 GluGlnSerPheHisLeuAla--LeuLeuArgGlnProGlyIleTyrLysProPheIleI 259
 |||||
 Db 767 GAGCAGAGCTTTCACCTTGGCCCTTGTGGGAAGCCCGGCGCATCTACAAGCCCTTCATCA 826
 QY 259 le 259
 Db 827 TC 828

RESULT 14
 BG717034
 LOCUS 777 bp mRNA linear EST 08-MAY-2001
 DEFINITION 602689093F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4821598 5',
 mRNA sequence.

ACCESSION BG717034
 VERSION BG717034.1 GI:13996221
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>;
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM10728 row: i column: 23
 High quality sequence stop: 722.

Location/Qualifiers
 1..777
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4821598"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_97"

FEATURES

source

/note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcag); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTTN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
 Pred. No.: 1.18e-106 Length: 777
 Score: 1233.00 Matches: 243
 Percent Similarity: 98.39% Conservatave: 1
 Best Local Similarity: 97.98% Mismatches: 3
 Query Match: 50.18% Indels: 2
 DB: 4 Gaps: 0

US-09-886-954A-1 (1-477) x BG717034 (1-777)

QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProGlySerAla 20
 Db ATGACGCCCGAGGACCAGAGAAACCCAGCCGCTTCGGGGCTTCCTGGCGGACGCGG 95
 QY 21 ProArgGlyArgValPheLeuAlaPheAlaAlaLeuGlyProLeuSerPhe 40
 Db CCCCAGCGCGCGCGCTTCCTCGCGGCTTCGGCGGCTTCGGCGCCACTCAGCTTC 155
 QY 41 GlyPheAlaLeuGlyTyrSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 Db 156 GGCTTCGGCTCGCTACAGCTCCCGCGCCATCCCTAGCTCGAGCGCCCGCCCG 215
 QY 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
 Db 216 GCCCGCGCGCTGACAGCCCGCGCTCCTGGTTCGGGGCTTCGTGACCCCTGGGTGC 275
 QY 81 AlaAlaGlyValLeuGlyTrpLeuValAspArgAlaGlyArgIleuSerLeu 100
 Db 276 CGCGCGCGGGAGTGTGGCGGCTGGCTGGTGGACCGCGCGCGGGCGAAGCTGAGCCTC 335
 QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 Db 336 TTGCTGTCTCCGTCCTCGTGGCGGCTTTGGCGTTCATCACCGCGGCCAGGACCTG 395
 QY 121 TrpMetLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 Db 396 TGGATGCTGTGGGGGGCGGCTTCCTCACCGGCTGGCTGGCGGTTCCTCCCTAGTG 455
 QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuGlySerCys 160
 Db 456 GCCCGCGGTTCACATCTCCGAAATCGGCTACCCAGCAGTCCGGGGGTTCCTGCTCT 515
 QY 161 ValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
 Db 516 GTGACGTAATGTGTGTGTGGCATCCCTCGGCTTACTGGCGGCTGGGTGGTGGAG 575
 QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 Db 576 TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 635
 QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
 Db 636 TTCATGCCGAGACCCCGCGCTTCCTGCTGACTCAGCACAGGCGCCAGGAGCCATGGCC 695
 QY 221 AlaLeuArgPheLeu-TrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaG 240
 Db 696 GCCCTGGGTTCCTGTTGGGGCTCCGAGCATGCTGGGAGACCCCGCCATC-GGAGCTGA 754
 QY 240 uGlnSerPheHisLeuAlaLeu 247
 Db 755 CCAGAGCTTCCACCTGGACCTG 776

RESULT 15

BI757409
 LOCUS 889 bp mRNA linear EST 25-SEP-2001
 DEFINITION 603029344F1 NIH_MGC_114 Homo sapiens CDNA clone IMAGE:5199778 5', mRNA sequence.
 ACCESSION BI757409.1 GI:15748987
 VERSION BI757409.1
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 889)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)
 cDNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM11500 row: 9 column: 11
 High quality sequence stop: 877.
 Location/Qualifiers

REFERENCE

1. 889
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5199778"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_114"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dr primed and directionally cloned [EcoRV site is destroyed upon cloning]. Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH_MGC Library."

FEATURES

source

ORIGIN

Alignment Scores:
 Pred. No.: 2.48e-106 Length: 889
 Score: 1230.50 Matches: 265
 Percent Similarity: 78.87% Conservatave: 0
 Best Local Similarity: 78.87% Mismatches: 4
 Query Match: 50.08% Indels: 69
 DB: 4 Gaps: 1

US-09-886-954A-1 (1-477) x BI757409 (1-889)

QY 143 ValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGln 162
 Db 38 GTCTACATCTCGAAATCGCTTACCAGCACTCCGGGGGTTCCTGGCTCGCTGGTGCAG 97
 QY 163 LeuMetValValValGlyIleLeuLeuAlaTyrIleuAlaGlyTrpValLeuGluTrpArg 182
 Db 98 CTAATGTCGTTCGTGGCATCTCCTGGCTTACCTGGCA----- 136
 QY 193 TrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPheMet 202
 Db 136 ----- 136
 QY 203 ProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaLeu 222
 Db 136 ----- 136

Qy 223 ArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProPheGlyAlaGluGlnSer 242
 Db 137 -----GAG 139

Qy 243 -PheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIleGlyValSe 262
 Db 140 CTTTACCTGGCCCTCTGGGAGCCCGGCATCTAACAGCCCTTCATCATCGGGGCTC 199

Qy 262 rLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGluThrI 282
 Db 200 CCTGATGGCTTCCAGCAGCTGTCCGGGGTCAACGCCGTCAATGTTCTATGACAGACCAT 259

Qy 282 ePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValValGlyValIleGI 302
 Db 260 CTTTGAAGAGGGCCAGTTCAAGGACAGCAGCCCTGGCTCGGTCTGGGTGTCATCCA 319

Qy 302 nValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeuLeuLe 322
 Db 320 GGTGCTGTTCAACAGCTGTGGGGCTCTCATCFGGACAGACAGGGCGGGGCTGCTCCT 379

Qy 322 uValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPheLysLe 342
 Db 380 GGTCTTGTCCAGGTGCTCATGCTGTTCCAGCAGAGTGCCTTCGGGGCTACTTCAAGCT 439

Qy 342 uThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSerAlaGI 362
 Db 440 GACCCA-GGTGGCCCTGGCAACTCTCCGACCTGGCCATCTCGGGCCCTGTCTCTGCACA 498

Qy 362 nProValAspAlaSerValGlyLeuAlaTrpLeuAlaValAlaGlySerMetCysLeuPheI 382
 Db 499 GCCTGTGTAGCCAGGCTGGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTTTCAT 558

Qy 382 eAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePhePr 402
 Db 559 CGCGGGCTTGGGGTGGGGTGGGGGCCCATCCCTGGCTCTCTCATGTTCAGAGATCTTCCC 618

Qy 402 oLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPh 422
 Db 619 TCTGCATGTCAAGGGCGTGGGACAGGCATCTCGCTCCCTCACCAACTGGCTCATGGCCCT 678

Qy 422 eLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTr 442
 Db 679 TCTGGACCAAGGAGTTCAGCAGCCTCATGGAGGTCCTCAGGCCCTATGGAGCCCTTCG 738

Qy 442 pLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluTh 462
 Db 739 GCTTGGCTCCG--TTTCTGATCTTCAGTGTCTTTTCACTTGTCTGTCTGTCTCCCTGAAC 796

Qy 462 rLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 Db 797 TAAAGGAAGACTCTGGAACAATAACAGCCCAATTTTGGGGGGCGA 842

Search completed: February 25, 2005, 21:31:27
 Job time : 3621 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 05:53:07 ; Search time 43 Seconds
(without alignments)
828.084 Million cell updates/sec

Title: US-09-886-954A-1
Perfect score: 2457
Sequence: 1 MTPDEPTEQPLGPPGSA.....CVPETKGLQITAHFEGR 477

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Issued Patents AA.*
1: /cgn2_6/prodata/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/iaa/6A COMB.pep.*
4: /cgn2_6/prodata/iaa/6B COMB.pep.*
5: /cgn2_6/prodata/iaa/PCTUS COMB.pep.*
6: /cgn2_6/prodata/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Table with columns: Result No., Score, Query Match, Length, DB ID, Description. Contains 27 rows of search results.

Table with columns: 28, 498.5, 20.3, 492, 5, PCT-US95-16126-3, Sequence 3, Appli. Contains 27 rows of sequence identifiers.

ALIGNMENTS

RESULT 1
US-09-949-016-9610
; Sequence 9610, Application US/99949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J, Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9610
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-9610

Query Match 49.5%; Score 1215.5; DB 4; Length 246;
Best Local Similarity 99.6%; Pred. No. 8.8e-107;
Matches 241; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 86 LGGWLVDRAGRKLSLLCSVPPVAGFAVITAAQDVWMLLGGRLITGLACGASLVAPVYI 145
DB 1 LGGWLVDRAGRKLSLLCSVPPVAGFAVITAAQDVWMLLGGRLITGLACGASLVAPVYI 60
QY 146 SIIAYPAVAGLGGSCVQLMVMVVGILLAYLAGWLEWRWLVLCVPPSLMLLLMCFMPET 205
DB 61 SIIAYPAVAGLGGSCVQLMVMVVGILLAYLAGWLEWRWLVLCVPPSLMLLLMCFMPET 120
QY 206 PRFLFTQHRQRQAMAALRELVGSEQGWEDPPIGAE-OSFHALLRQPGYKPFIIIGVSLM 264
DB 121 PRFLFTQHRQRQAMAALRELVGSEQGWEDPPIGAEQQSFHALLRQPGYKPFIIIGVSLM 180
QY 265 AFQQLSGVNVAFYAEITFEAKFKDSSLASVVVGVIVLFTVAVALIWDVAGRRLLLVL 324
DB 181 AFQQLSGVNVAFYAEITFEAKFKDSSLASVVVGVIVLFTVAVALIWDVAGRRLLLVL 240
QY 325 SG 326
DB 241 SG 242

QY 302 QVLTAAALIMDRAGRRLLVLSGVVVFSTSAFGAYFKLTQGGFGNSHVAISAPVSA 361
 Db 348 FMPATFIAVTVDKAGRKPALKIGFSVMALGTLVLG--YCLMQFDNGT----- 393
 QY 362 QPVDASVGLAWLAVGSMCLFIAGFAVGVGPIWLLMSEIFPLHVKGVATGICVLINWLMMA 421
 Db 394 ----ASSGLSWLSVGNMCIAGYAMAAPVVMWILCSEIQPLKRCDFGTCSTTTNWSN 449
 QY 422 FLVTKFSSIMEVLRPYGAFWLASAFICFSLVFTLFCVPETKGTLEQI 470
 Db 450 MIIGATFLILLDAIGAAGTFWLYTALNVAFIGITFWLIPETKRVILDEHI 498

RESULT 4
 US-09-489-039A-11933
 ; Sequence 11933, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489, 039A
 ; PRIORITY FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 11933
 ; LENGTH: 476
 ; TYPE: PRT
 ; ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-11933

Query Match 23.5%; Score 578.5; DB 4; Length 476;
 Best Local Similarity 31.2%; Pred. No. 3.8e-46;
 Matches 149; Conservative 84; Mismatches 180; Indels 65; Gaps 10;

QY 27 FLAAPAAALGSLGFGALGYSIPAIPSLQRAAPPAPRLDDAAAFWFGAVVTLGAAGGYL 86
 Db 28 FFCFLAALAGLLFDLDGVIAGLPIANEF----QISAHQEWVSSMFGAAVAVG 93
 QY 87 GGLVDRAGKLSLLCSVFPVAGFAVITAAQDVMMLLGGRLTLGACGVASLAVAPVYIS 146
 Db 84 SGWLSFKLRKSLMIGAILFVAGSLFSAAPNPVEILLVSRVLLGLAVGASVYAPLYLS 143
 QY 147 EIAYPAVRLGSCVQLMVVVGILLAYLAGWLE----WRWLAVLCCVPPSLMMLMCFM 202
 Db 144 EIAPEKIRGSMISMYQLMITIGILGAYLSDTAFSYSAGWRWMLGVIIIPAVLLIGVIFL 203
 QY 203 PETPFLLTQHRQEAAMAL-----RFLMGSEOGWEDPPPIGAQSFH 244
 Db 204 PDSRPFPAKRFRVDAEVLRLRDLTSAEKRELDRELSKVKQSGW----- 251
 QY 245 LALLRQFGIYKPI-IGVSLMAFQQLSGVNAVMYAETIFEEAKFKDSS---LASVYVG 300
 Db 252 -SLFKNSNFRRAVFLGILLQVMQQTGMVNIYVYAPKIFELAGYANTTEQMGTIVGL 310
 QY 301 IQVLTAAALIMDRAGRRLLVLSGVVVFSTSAFGAYFKLTQGGFGNSHVAISAPVS 360
 Db 311 TNVLTATFIAIGLVDWRGRKPTLLGFIW-----AAG-----MGVLTMMHIGHSSTA 359
 QY 361 AQPVDASVGLAWLAVGSMCLFIAGFAVGVGPIWLLMSEIFPLHVKGVATGICVLINWLM 420
 Db 360 -----QYIATVLMMLFIVGFAMWAGLWVLCSEIQPLKGRDFGTCSTANWIA 409
 QY 421 AFLVTKFSSIMEVLRPYGAFWLASAFICFSLVFTLFCVPETKGTLEQITAH-FEGR 477
 Db 410 NMIVGATFLTNLSGNSANTFWYGGNLVLFILLTLLMIPETKNSLSIERNLMAQR 467

RESULT 5
 US-09-031-392-5

; Sequence 5, Application US/09031392
 ; Patent No. 5942398
 ; GENERAL INFORMATION:
 ; APPLICANT: Tartaglia, Louis A.
 ; APPLICANT: Weng, Xun
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES
 ; TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows95
 ; SOFTWARE: FastSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/031,392
 ; FILING DATE: 26-FEB-1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Meiklejohn, Ph.D., Anita L.
 ; REGISTRATION NUMBER: 35,283
 ; REFERENCE/DOCKET NUMBER: 07334/072001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617/542-5070
 ; TELEFAX: 617/542-8906
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 494 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-09-031-392-5

Query Match 23.4%; Score 575.5; DB 2; Length 494;
 Best Local Similarity 30.7%; Pred. No. 7.6e-46;
 Matches 150; Conservative 94; Mismatches 179; Indels 65; Gaps 10;

QY 26 VFLAFAAALGSLGFGALGYSIPAIPSL-----QRAAPPAPRLDDAAAFWFGAV 75
 Db 10 LIFALSIATIGSFQGYNTGVINAPEAIKDFLNYLLEERSETPPSVLLTSLWSLSVAI 69
 QY 76 VTGAAAGVGLGWLVDRAGRKLSLLCSVFPVAGFAVI---TAAQDVMMLLGGRLTLGL 132
 Db 70 FSVGGMIGSFSVGLFVNRFGRENSMLVNLAIAGCLMGFCFKIAESVEMLIILGRIIGL 129
 QY 133 ACGVASLAVPVYISETAYPAVRLGSCVQLMVVVGILLAYLAGWLE----EWRWLAVLG 188
 Db 130 FCGLCTGFVPMYIGEISPTALRGAFTLNLQGLGIVIGILVAQIFGLKVIILGTELDWPELLG 189
 QY 189 --CVPSLMLLMLCMFPEPPPELLTQHRQE-AMAAALRFLMGSEOGWEDPPPIGAQSFH- 244
 Db 190 FILLPAIQCAALPCPESPPELLNKRKEEKAKELQRLWGTEDVAQIQEMKDESMRM 249
 QY 245 -----LALLRQFGIYKPIIGVSLMAFQQLSGVNAVMYAETIFEEAKFKDSSLASV 296
 Db 250 SQEKQVTVLELFRAPNYRQPIIISIMLQLSQQLSGINAVFYSTGIFKDGAVQEPVYATI 309
 QY 297 VVGVQLVFTAAALIMDRAGRRLLVLSGVVVFSTSAFGAYFKLTQGGFGNSHVAIS 356
 Db 310 GAGVNTIFTVVSFLVERAGRRLHLI-----GLGMAFCSIL 348
 QY 357 APVSAQVDSVGLAWLAVGSMCLFIAGFAVGVGPIWLLMSEIFPLHVKGVATGICVL 416
 Db 349 MTSILLKDNYSWMSFICIGAILVFAVFEIGFPIWFLVIAELFQGGPRPAAAVAGCS 408
 QY 417 NWLMAFLVTKFSSIMEVLRPYGAFWLASAFICFSLVFTLFCVPETKGTLEQ 469


```

Db 70 FSVGGMIGSFVGLFVNRFRGRNSMLVNLALTAGCLMGFCRKAESVEMLLIQLRLLI 129
QY 133 ACGVASLAVPVYISETAYPAVRGLLSCVQLMNVVVGILLAYLAGWVL---EWRWLAVLG 188
Db 130 FCGLCTGFVPMYIGESPTALRAFGFTLNQLGIVIGILLVAQIFGLKVLGTEDLWPLLLG 189
QY 189 --CVPSLMLLMLCFMPEPRFLLTQHRROE-AMAALRFLWMSQGWEDDPPIGAQSFH- 244
Db 190 FTLLPAILQCAALPFCPEPRFLLINRKEBEKAKEILQRLWGTEDVAQDIQEMKDESMRM 249
QY 245 -----LALLRQPGIYKPFIIIGVSLMAFQQLSGVNVVWFYAETIPEERAKFKDSSLASV 296
Db 250 SOBKQVTVLELFRVSSYRQPIIISIVLQLSQQLSGINAVFYSTGIFKDGAVQEPYATI 309
QY 297 VVGVIQVLFVAVAAALMDRAGRLLLVLSGVVWVFTSARFAYFKLTQGGPGRSSHVAIS 356
Db 310 GAGVWNTIFVSVFLVERAGRRLHMI-----GLGMAFCSTL 348
QY 357 APVSAQPDVAVGLAWLAVGSMCLFIAGFVAVGMPPIWLLMSRIFPLHVKGVATGICVLT 416
Db 349 MTISLLKDNYSNWSFCIGAILLVFAVFEIGPGPIWFIWVIAELFSGQPRPAAVAVGCS 408
QY 417 NMLMAFLVTKFSSLMELRYPYGAFWL-ASAFICFSV-----LFTLFCVPETKGTLEQ 469
Db 409 NMTSNEFLV-----GLLFPSSATFYLGAVFVFTVFLVIFWVFTFFKVPETRGRTFEE 460
QY 470 ITRAFEGR 477
Db 461 ITRAFEGQ 468

```

```

RESULT 8
US-10-146-704-3
; Sequence 3 Application US/10146704
; Patent No. 6828421
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Incorporated
; APPLICANT: Sugiyama, Janice
; TITLE OF INVENTION: VAMP-ASSOCIATED PROTEIN A-INTERACTING PROTEINS AND USE THEREOF
; FILE REFERENCE: 1669 01
; CURRENT APPLICATION NUMBER: US/10/146, 704
; CURRENT FILING DATE: 2004-05-14
; PRIOR APPLICATION NUMBER: US 60/291,730
; PRIOR FILING DATE: 2001-05-17
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-704-3

```

```

Query Match 23.2%; Score 570.5; DB 4; Length 496;
Best Local Similarity 31.6%; Pred. No. 2.3e-45;
Matches 152; Conservative 88; Mismatches 190; Indels 51; Gaps 9;
QY 26 VFLAAFAALGPLSFGFALGYSSPAIFSLQR-----RAPPAPRLDDAAASWFGAV 75
Db 10 LIFAITVATIGSFQGYNTGVINAPKEIKKEFINKTLTKGNAPPSEVLLTSLMSLSVAI 69
QY 76 VTLLGAAAGGVLGGWLDVDRGKLSLLLCSPVFFVAG---FAVITAAQDVMWLLGGRLLTGL 132
Db 70 FSVGGMIGSFVGLFVNRFRGRNSMLVNLALTAGCLMGFCRKAESVEMLLIQLRLLI 129
QY 133 ACGVASLAVPVYISETAYPAVRGLLSCVQLMNVVVGILLAYLAGWVL---EWRWLAVLG 188
Db 130 FCGLCTGFVPMYIGESPTALRAFGFTLNQLGIVIGILLVAQIFGLKVLGTEDLWPLLLG 189
QY 189 --CVPSLMLLMLCFMPEPRFLLTQHRROE-AMAALRFLWMSQGWEDDPPIGAQSFH- 244
Db 190 FTLLPAILQCAALPFCPEPRFLLINRKEBEKAKEILQRLWGTEDVAQDIQEMKDESMRM 249

```

```

QY 245 -----LALLRQPGIYKPFIIIGVSLMAFQQLSGVNVVWFYAETIPEERAKFKDSSLASV 296
Db 250 SOBKQVTVLELFRVSSYRQPIIISIVLQLSQQLSGINAVFYSTGIFKDGAVQEPYATI 309
QY 297 VVGVIQVLFVAVAAALMDRAGRLLLVLSGVVWVFTSARFAYFKLTQGGPGRSSHVAIS 356
Db 310 GAGVWNTIFVSVFLVERAGRRLHMI-----GLGMAFCSTL 348
QY 357 APVSAQPDVAVGLAWLAVGSMCLFIAGFVAVGMPPIWLLMSRIFPLHVKGVATGICVLT 416
Db 349 MTISLLKDNYSNWSFCIGAILLVFAVFEIGPGPIWFIWVIAELFSGQPRPAAVAVGCS 408
QY 417 NMLMAFLVTKFSSLMELRYPYGAFWL-ASAFICFSV-----LFTLFCVPETKGTLEQ 469
Db 409 NMTSNEFLV-----GLLFPSSATFYLGAVFVFTVFLVIFWVFTFFKVPETRGRTFEE 460
QY 477 R 477
Db 468 Q 468

```

```

RESULT 9
US-09-031-392-10
; Sequence 10 Application US/09031392
; Patent No. 5942398
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESS: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,392
; FILING DATE: 26-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 493 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-031-392-10

```

```

Query Match 23.1%; Score 568; DB 2; Length 493;
Best Local Similarity 32.2%; Pred. No. 3.9e-45;
Matches 159; Conservative 87; Mismatches 160; Indels 88; Gaps 14;
QY 28 LAFAAALGPLSFGFALGYSSPAIFSLQRAPAPRLDDAAASWFG-----AVVTLGA- 80
Db 16 LAVLIAALGSFQGYNLGVIN-----APQVIEAFYETLWLRGXEXSPVPTLTL 65
QY 81 -----AAGVILGWLV-----DRAGKLSLLLCSPVFFVAG---FAVITAAQDVMWLLG 125
Db 66 WLSVSIFAVGGMIGSFVGLVGNRIGRICKXKAMLVNVLAIAGLLMGLAKKXAXSFFMLIL 125

```

Query Match 23.1%; Score 568; DB 3; Length 493;
 Best Local Similarity 32.2%; Pred. No. 3 9e-45;
 Matches 159; Conservative 87; Mismatches 160; Indels 88; Gaps 14;

Qy 28 LAFAAALGALGSLGFGALGYSSPAIPSLQRAAPAPRDLDDAAASWFG-----AVVTLGA- 80
 Db 16 LAVLAAALGSGFYGYNLGVN-----APQKVEAFYETWLGXGEXPSVPTLL 65

Qy 81 -----AAGVLLGGWLV-----DRAGKLSLLCSVPFVAG---FAVITTAQDVMWLLG 125
 Db 66 WLSLVSIVFAYGGMIGSFLVXGIGNRGLGRKXKAMLVNVLATAGLLMLGLAKKXKXSFEMLIL 125

Qy 126 GRLLTGLACGVASIVAPVISEIAYPAVRGLLGSQVLMVVVGGILLAYLAGWLE----- 180
 Db 126 GRFIIIGLYCGLSSGVVPMYVGEISPTALRGALGTINQLGIVIGLIIAQVLG--LDSLLGN 183

Qy 181 ---WRWLAVLGCVPPSLMILLMCMFETPRFLLTOHRRQEMA--ALRFLWG----- 227
 Db 184 ESLWPLLLGLTGVPAQLQLLLLPFCPESPRYLLI-NKNEERAKKALQRLRGTDVDSQEV 242

Qy 228 SEQGWEDPPIGAQSFH-LALLRQPGIYKPFIIIGVSLMAFQOLSGVNAVVFYAEITFEEA 286
 Db 243 AEMKDESRXKXSEKXSVLEFRXRYRQPVIIAIVLQLSQLSGINAVFYYSIFEKA 302

Qy 287 KFKDSSLASVVVGVIOVLFQVAVLIMDRAGRLLLVLSGVVMVVFSTSAFGAYFKLTGG 346
 Db 303 GVGQPVYATIGAGVNVTVFVVSFVVERAGRRTLLH-----G 341

Qy 347 PGNSSHVAISAPVSAQPDVAVGLWLAAGSMCLFIAGFVAGWGPPIWLLMSEIFPLHVX 406
 Db 342 LGMAGCAVMTIALALDQVPMYSVIVAIFFGVAFVFEVGGPIPFVIAELFSQGR 401

Qy 407 GVATGICVLTWMLMAFLVTKFSSLMELVLRPYGAFWLASAFICFISVL-----FTLFCVP 460
 Db 402 PAAIYAVAGFSNWTNFVGLLFOYIAELGPI-----VFIVFAVLLLLFFIFTLKVP 454

Qy 461 ETGKLEQITAHF 474
 Db 455 ETGRTFDEIAAAF 468

Qy 126 GRLLTGLACGVASIVAPVISEIAYPAVRGLLGSQVLMVVVGGILLAYLAGWLE----- 180
 Db 126 GRFIIIGLYCGLSSGVVPMYVGEISPTALRGALGTINQLGIVIGLIIAQVLG--LDSLLGN 183

Qy 181 ---WRWLAVLGCVPPSLMILLMCMFETPRFLLTOHRRQEMA--ALRFLWG----- 227
 Db 184 ESLWPLLLGLTGVPAQLQLLLLPFCPESPRYLLI-NKNEERAKKALQRLRGTDVDSQEV 242

Qy 228 SEQGWEDPPIGAQSFH-LALLRQPGIYKPFIIIGVSLMAFQOLSGVNAVVFYAEITFEEA 286
 Db 243 AEMKDESRXKXSEKXSVLEFRXRYRQPVIIAIVLQLSQLSGINAVFYYSIFEKA 302

Qy 287 KFKDSSLASVVVGVIOVLFQVAVLIMDRAGRLLLVLSGVVMVVFSTSAFGAYFKLTGG 346
 Db 303 GVGQPVYATIGAGVNVTVFVVSFVVERAGRRTLLH-----G 341

Qy 347 PGNSSHVAISAPVSAQPDVAVGLWLAAGSMCLFIAGFVAGWGPPIWLLMSEIFPLHVX 406
 Db 342 LGMAGCAVMTIALALDQVPMYSVIVAIFFGVAFVFEVGGPIPFVIAELFSQGR 401

Qy 407 GVATGICVLTWMLMAFLVTKFSSLMELVLRPYGAFWLASAFICFISVL-----FTLFCVP 460
 Db 402 PAAIYAVAGFSNWTNFVGLLFOYIAELGPI-----VFIVFAVLLLLFFIFTLKVP 454

Qy 461 ETGKLEQITAHF 474
 Db 455 ETGRTFDEIAAAF 468

RESULT 10
 US-09-299-549-10
 ; Sequence 10, Application US/09299549
 ; Patent No. 6136547
 ; GENERAL INFORMATION:
 ; APPLICANT: Tartaglia, Louis A.
 ; APPLICANT: Meng, Xun
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
 ; TITLE OF INVENTION: GLUTEX AND USES THEREOF
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Fish & Richardson P.C.
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows95
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/299,549
 ; FILING DATE: 26-APR-1999
 ; PRIORITY INFORMATION:
 ; APPLICATION NUMBER: 09/031,392
 ; FILING DATE: 26-FEB-1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Meiklejohn, Ph.D., Anita L.
 ; REGISTRATION NUMBER: 35,283
 ; REFERENCE/DOCKET NUMBER: 07334/072002
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617/542-5070
 ; TELEFAX: 617/542-8906
 ; TELEX: 200154
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 493 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-299-549-10

RESULT 11
 US-09-610-417-10
 ; Sequence 10, Application US/09610417
 ; Patent No. 6346374
 ; GENERAL INFORMATION:
 ; APPLICANT: Tartaglia, Louis A.
 ; APPLICANT: Meng, Xun
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
 ; TITLE OF INVENTION: GLUTEX AND USES THEREOF
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Fish & Richardson P.C.
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows95
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/610,417
 ; FILING DATE: 05-Jul-2000
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/299,549
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Meiklejohn, Ph.D., Anita L.
 ; REGISTRATION NUMBER: 35,283

```

; REFERENCE/DOCKET NUMBER: 07334/072002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 493 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-610-417-10

```

```

Query Match 23.1%; Score 568; DB 3; Length 493;
Best Local Similarity 32.2%; Pred.No. 3.9e-45;
Matches 159; Conservative 87; Mismatches 160; Indels 88; Gaps 14;
QY 28 LAFAAALGFLGFGALGYSFAIPSLORAAPPAPRLDAAAASFQ-----AVVTLGA- 80
DB 16 LAVLIAALGSGFQYGNLVGIN-----APQVIEAFYEWLGRXGEXPSVPTLTL 65
QY 81 -j-----AAGVIGGLV-----DRAGRKLSLLCSVPFVAG---PAVITAAQDVWMLL 125
DB 66 WSLSVSIFAVGGMIGSLVXIGNRLGRKXKAMLVNNVLAIAAGLLMGLAKXAXSFEMLL 125
QY 126 GRLLTGLAGVASLAVPVISETAYPAVRGLLGSVCQLMVMVVGILLAYLAGWLE----- 180
DB 126 GRPIIIGLYCGLSSGVVPMVXGEISPTALRGALGTLLNQLGIVIGILLIAQVIG--LDSLLGN 183
QY 181 ---WRWLAVLGCVPPSMLLMLLGFMPETPRFLITQHRROEAMA--ALREFLWG----- 227
DB 184 ESLWPLLGLGTVPALLQLLFLFCPESEPRYLLI-KNEEAREAKALORLRTADVYSQEV 242
QY 228 SEGGWEDPPIGAPQSPH-LALLRQPGIYKPFIIIGVSLMAFQQLSGVNAVVFYAEIIFEEA 286
DB 243 AEMKDSRAXMKSEKXVLELFRSRYRQPVIIAIVLQSLQSLGGINAVFYIYSTSIFEXA 302
QY 287 KFDSSLASVVVQIVLFTAVNAALIMDRAGRLLLVLSGVVMVFTSFAFGYFKLTOG 346
DB 303 GVGQPYATIGAGVNVTVFVSVFVERAGRRLTL-----G 341
QY 347 PGNSSHVVAISAPVSAQPVASVGLAWLAVGSMCLFIAGFAVGMPIPLMLMSEIIFPLHVK 406
DB 342 LGMACCAVLMITALLLQVPMMSVSIVAIFGVFAFEVGGPIPWFIWVIAELFSQGR 401
QY 407 GVATGICVLTNMLMAFLYKESLMEVLRPYGAFWLASAFICFISVLE-----FTLFCVP 460
DB 402 PAAIAVAGFSNMTSNEIFVIGLLFOYIAELLGPY-----VFIVFAVLLLLFFITFTELKVP 454
QY 461 ETGKKTLEOITAHF 474
DB 455 ETGKRTFDEIAAAF 468

```

```

RESULT 12
US-09-679-686B-23
; Sequence 23, Application US/09679686B
; Patent No. 6624343
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: HEXOSE CARRIER PROTEINS
; FILE REFERENCE: BB1160 US NA
; CURRENT APPLICATION NUMBER: US/09/679,686B
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: 60/081,131
; PRIOR FILING DATE: 1998-04-09
; NUMBER OF SEQ ID NOS: 24
; ORGANSIM: Microsoft Office 97
; SEQ ID NO: 18
; LENGTH: 517
; TYPE: PR
; ORGANSIM: Triticum aestivum
; US-09-679-686B-18

```

```

; SOFTWARE: Microsoft Office 97
; SEQ ID NO 23
; LENGTH: 518
; TYPE: PR
; ORGANSIM: Medicago truncatula
; US-09-679-686B-23
Query Match 22.3%; Score 548.5; DB 4; Length 518;
Best Local Similarity 30.4%; Pred.No. 2.9e-43;
Matches 151; Conservative 88; Mismatches 190; Indels 67; Gaps 16;
QY 16 PGGSAAPRRRVFLAFAAALGPIGFGALGYSSPAI---PSLQRAAPPAPR----- 63
DB 16 PGNLTP---FVITITCIVAAAGGLIFGYDIGISGVVTSMDFLKKFFPAPVYRKNKDKSTN 72
QY 64 ----LDDAAASWFGAVVTGLGAAAGVVGWLVDRGRKLSLLCSVPFVAGFAVITAAQD 119
DB 73 QYCYDQSOTLWFTSLLYLAALLSLVASTITRRFRKLSMLFGGLLFLVGLALINGFANH 132
QY 120 VMMLGGRLITGLACGVASLAVPVISETAYPAVRGLLGSVCQLMVMVVGILLAYLAGWVL 179
DB 133 VMMLIVGRILLGFGIFANQPVPLVSEWAPYKRGALNIGFQLSITIGILVANVLYPF 192
QY 180 E-----WRW-LAVLGCVPPSMLL-LMCFMPEPRFLITQHRROEAMAALREFLW 229
DB 193 AKIKGWGWRSLGGMVPAIITIGSLVLPPTNPSMIERGDGRDGAQAKRIRGIEDVD 252
QY 230 QGWEDPPIGAEQSFLA-----LLRQPGIYKPFIIIGVSLMAFQQLSGVNAVVFYAE 283
DB 253 EEFNDLVAASEASMQVENPWRNLQRK---YRQPTMVAVLIPFQPTGINVIMFYAPVLF 310
QY 284 BEAKFKD--SLSASVVVGVIVLFTAVNAALIMDRAGRLLLVLSGVVM-----VF 337
DB 311 NSIGFKDDASLMSAVITGVVNVVAVCSYIGYDKVGRRALFEGGQMLICQVAVAAAG 370
QY 338 AYPKLTQGGPNSHSHVAISAPVSAQPVASVGLAWLA---VGSMLCFIAGFAVGMPIPW 394
DB 371 AXFG-TSGNPNLPE-----WYAVVVLFCIYVAGFAMSWGPLGW 410
QY 395 LMSIIFPLHVKAVTIGVLTNMLMAFLYKESLMEVLRPYGAFWLASAFICFISVLF 454
DB 411 LVPSEIIFLEIRSAASQVNVSYNMLFTFLVAQVFLMLCHMK-FGLFLFFAFVFLVMSY 469
QY 455 TLFVCPETKGTLEOI 470
DB 470 VFLLPETHKGIPEEM 485

```

```

RESULT 13
US-09-679-686B-18
; Sequence 18, Application US/09679686B
; Patent No. 6624343
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: HEXOSE CARRIER PROTEINS
; FILE REFERENCE: BB1160 US NA
; CURRENT APPLICATION NUMBER: US/09/679,686B
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: 60/081,131
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: PCT/US99/07561
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO: 18
; LENGTH: 517
; TYPE: PR
; ORGANSIM: Triticum aestivum
; US-09-679-686B-18

```

Query Match 22.1%; Score 542.5; DB 4; Length 517;
 Best Local Similarity 29.9%; Pred. No. 1.1e-42;
 Matches 150; Conservative 88; Mismatches 202; Indels 61; Gaps 14;

QY 17 GCSAPRR---VFLAFAAALGPLSFGFALGSSPAI---PSLQRAAPP-----61
 DB 11 GCKDPGRLLTFFVFCVVAATGGLIFGYDIGISGVTSMNPLKFKFPEVYDKQKMG 70
 QY 62 ---PRLDDAASWFGAVTLCGAAGVGLGGVLDVDRAGKLSLLCSVPPFVAGAVITAA 117
 DB 71 ASQCKYDNLQQLQTTSSLYLALVSSFFAAVTRVVGKRWKSMFTGGLTFLIGAALNGAA 130
 QY 118 QDVWMLLGRLLTGLACGASLAVPVYISEIAYPAVRGLLSCVQLMVMVVGILLAYL--- 174
 DB 131 ENIAMLIVGRILLGVGVGFANGVVPVYLSMAPARLGMNLGFMQMITIGILAAALNY 190
 QY 175 -----AGVLEWRMLAVLGCVPPLSMLLLMCFMPTFRFLLTQHRROBMAALRFLMGS 228
 DB 191 DTNKIKAGY--GWRISLAIAAVALPAGIITLGSFFLPTDPTNSLIERGHPAARRMLNIRGS 248
 QY 229 E---QGWEDPPIGAEQSHLALLROP-----GIYKP-FIIGVSLMAFQQLSGVNAVWF 277
 DB 249 DVDISEYADLVVASEES---KLVQHPWRNIIQKRPQLTWAIMIPFQQUTGINVIMF 305
 QY 278 YAEITPEEAKFK-DSSLAS-VVVGVIQVLTVAALIMDRAGRRLLLVLSGVVWVFTSA 335
 DB 306 YAPVLFETLGFKGDASLMSAVITGLVNVFPATLVSFTVDRLGRKLFLOGGTQMLLSQLV 365
 QY 336 FGAYFKLTCGGGNSHVAISAPVSAQVDSVGLAWLAVGSMCLFIAGFVAGWGPWPWL 395
 DB 366 VGLTIAVFKFTSG-----VGENP-----KGYAAAVALFICLYVAGFAWSWGPLGLWL 411
 QY 396 LMSLFPPLHVKGVATGICVLTNMLMAFLVTKFSSLMELVRYGAPFLASAFICFSVLFT 455
 DB 412 VSEIFPLEIRPAGOSINVSNNLFTFVIAQAFLITWCHKM-FGLYFFAGVVMVIMTVEI 470
 QY 456 LFCVPETKTKLEQITAHFEG 476
 DB 471 ALFLPETKNVPIEMVLMVKG 491

RESULT 14
 US-09-031-392-6
 ; Sequence 6, Application US/09031392
 ; Patent No. 5942398
 ; GENERAL INFORMATION:
 ; APPLICANT: Tartaglia, Louis A.
 ; APPLICANT: Weng, Xun
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows95
 ; SOFTWARE: FastSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/031,392
 ; FILING DATE: 26-FEB-1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Meiklejohn, Ph.D., Anita L.
 ; REGISTRATION NUMBER: 35,283
 ; REFERENCE/DOCKET NUMBER: 07334/072001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617/542-5070
 ; TELEFAX: 617/542-8906

TELEX: 200154
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 509 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-09-031-392-6

Query Match 21.8%; Score 536; DB 2; Length 509;
 Best Local Similarity 30.5%; Pred. No. 4.3e-42;
 Matches 159; Conservative 86; Mismatches 172; Indels 104; Gaps 16;

QY 13 IGPGGSPAPRR---RVFLAFAAALGPLSFGFA-----LGVSSPA 50
 DB 8 IGSEDEPQORVYVTVLAVFSAVLGSLQFGYINAVQKVIQEQSYNETWLRQGP 67
 QY 51 IFSLQRAAPPAPRLDDAAASWFGAVVTLGA-----AAGVGLGWLW---DRGRKLSL 100
 DB 68 GPS---SIPP-----GTLTTLWALVSAIFSVGGMISFLIGISQWLGKRAM 112
 QY 101 LLCVPPFVAG---FAVITAAQDVWMLLGGRLTGLACGASLAVPVYISEIAYPAVRGLL 157
 DB 113 LVNVVAVLGGSLMGLANAAASYEMLIILGRFLIGAYSGLTSLVPMVYVGEIAPTHLRGAL 172
 QY 158 GSCVQLMVMVVGILLAYLAGWLE-----WRWLVAVLGCVPPLSMLLLMCFMPTFRFL 209
 DB 173 GTLQQLAVIGLLIAQVVG--LESLLGTASLWPLLLGLTVPALLQLVLLFCPPSPRYL 230
 QY 210 -LTQHRROBMAALRFL--WGSEOG---WEDPPIGAEQSHLALLRQPGI---YKPFII 259
 DB 231 YIQNLGEPARKSLKRLTGWADVSLVAELKDKRKLRRERPLSLQLLGSRTHRQPLII 290
 QY 260 GYSLMAFQQLSGVNAVMPYAEITPEEAKFKDSSLASVVVVIQVLTVAALIMDRAGR 319
 DB 291 AVVLQLSQQLSGINAVFYSTSIPTAGVGGPAPVATIGAVGVVTVTLVSVLLVERAGR 350
 QY 320 LLLVLSGVVWVFTSAFGAYFKLTCGGGNSHVAISAPVSAQVDSVGLAWLAVGSMC 379
 DB 351 TLHLL-----GLAGWCGCALMTVALLLRLERVPMYSVIVAI 389
 QY 380 LFIAGFVAGWGPWPPLMSEIFPLHVKGVATGICVLTNMLMAFLVTKFSSLMELVRYG 439
 DB 390 GFVAFFEIGPGPIWFIVAELFSQGPRAAMAVAGFSNWTNFIIGMFGVVAEAMGPFY- 448
 QY 440 AFWLASAFICFSVL-----FTLFCVPETKTKLEQITAHF 474
 DB 449 -----VFLFAVLLGLGFFIFFLRVPETGRTRTFDQISA 483

RESULT 15
 US-09-299-549-6
 ; Sequence 6, Application US/09299549
 ; Patent No. 6136547
 ; GENERAL INFORMATION:
 ; APPLICANT: Tartaglia, Louis A.
 ; APPLICANT: Weng, Xun
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows95
 ; SOFTWARE: FastSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/299,549
 ; FILING DATE: 26-FEB-1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Meiklejohn, Ph.D., Anita L.
 ; REGISTRATION NUMBER: 35,283
 ; REFERENCE/DOCKET NUMBER: 07334/072001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617/542-5070
 ; TELEFAX: 617/542-8906

```

; APPLICATION NUMBER: US/09/299,549
; FILING DATE: 26-APR-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/031,392
; FILING DATE: 26-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 509 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-299-549-6

```

```

Query Match          21.8%; Score 536; DB 3; Length 509;
Best Local Similarity 30.5%; Pred.No. 4.3e-42;
Matches 159; Conservative 86; Mismatches 172; Indels 104; Gaps 16;

```

```

QY 13 LGPPGSAAPRGR---RVFLAAFAAALGPLSFGFA-----LGYSSPA 50
; : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 8 IGSEDEPPQORVTGLVAVSFLVSLQFGYNGVINAPQKVIQSYNETWLGKQPE 67
; : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 51 IPSLQRAAPAPRLDDRAASWFGAVVTLGA-----AAGVLCGWLV----DRAGRKLSL 100
; : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 68 GPS---SIPP-----GTLTLMALSVAIFSVGGMISSFLIGIISQWLGRKRAM 112
; : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 101 LLCVSPFVAG---FAVITRAQDVWMLLGGRLTGLACGVASLVAPVYVISEIAYPAVRGLL 157
; : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 113 LVNVLAVLGGSLMGLNAAASYEMLLGRFLIGAYSLTSGLVPMVGEIAPTHLRGAL 172
; : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 158 GSCVQLMVMVVGILLAVLAGHVLE-----WRWLVLCVPPSPMLLLMCFMPTPREL 209
; : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 173 GTLNQLAIVIGILIAQVLG--LESLLGTASLWPLLGLTVLPALLQLVLLPFCPEPRYL 230
; : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 210 -LQHRROEAMALRFI--WGSEGG---WEDPPIGAQSFLHALLRQPGI----YAPFII 259
; : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 231 YIIQNLQEGPARKSLKRLTGWADVSGVLAELKDKRKLKLERERPLSLQLLGSRTHRQPLII 290
; : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 260 GVSLMAFOQLSGVNAVWVYAEITIFERAKFKDSSLASVWVGIQVLFVAVAAALIMDRAGR 319
; : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 291 AVVLQLSQQLSGNAVYFYSTIFETAGVQPAYATIGAGVNTVFTLVSVLLVERAGR 350
; : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 320 LLLVLSGVVMVFTSAFGAYFKLTOGGPGNSHVAISAPVSAQPVDAVGLAWLAVGSMC 379
; : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 351 TLHL-----GLAGMCGCAILMTVALLLRRVPMYSVIVAIF 389
; : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 380 LFTAGFVNGPPIPLWLMSEIFPLHKVATGCVLTNMLMAFLVTKFSSLMVLRPYG 439
; : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 390 GFVAFFEIGPPIPWFIABLFSGQPRPAAMAVAFSNWTSNFIIGMGFYVAEAMGPY- 448
; : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 440 AFWLASAFCIFFVL-----FTLFCVPETKGTLEQITAHF 474
; : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 449 -----VFLFVALLLGGFFIFFLRVPETRGRFTDQISAAP 483

```

```

Search completed: February 24, 2005, 06:05:26
Job time : 45 secs

```

THIS PAGE BLANK (USPTO)